

Modeling the Temporal Behavior of Human Color Vision for Lighting Applications

Citation for published version (APA):

Kong, X. (2021). *Modeling the Temporal Behavior of Human Color Vision for Lighting Applications*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Industrial Engineering and Innovation Sciences]. Technische Universiteit Eindhoven.

Document status and date:

Published: 05/07/2021

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

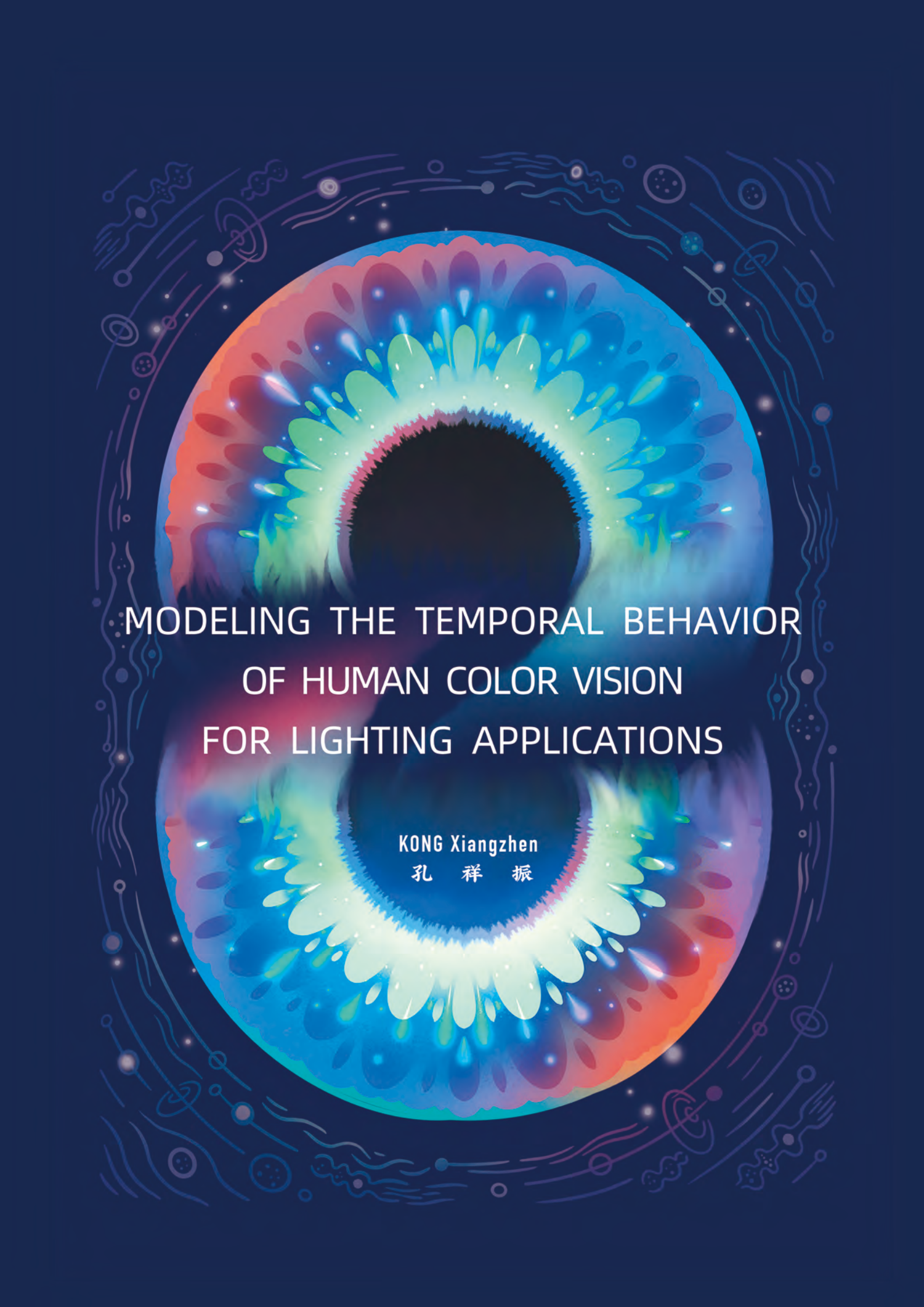
www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.



MODELING THE TEMPORAL BEHAVIOR
OF HUMAN COLOR VISION
FOR LIGHTING APPLICATIONS

KONG Xiangzhen
孔 祥 振

Modeling the Temporal Behavior of Human Color
Vision for Lighting Applications

Xiangzhen Kong



The research presented in this dissertation was mainly conducted at the Eindhoven University of Technology and partially at Rochester Institute of Technology and The Hong Kong Polytechnic University. The work described in this dissertation was performed within the framework of the strategic joint research program on Intelligent Lighting between Signify and TU/e. The research was mainly funded by the China Scholarship Council (CSC), and partially supported by Eindhoven University of Technology, Wuhan University of Technology and The Hong Kong Polytechnic University.

Copyright© Xiangzhen Kong (孔祥振), 2021

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the copyright holder.

A catalogue record is available from the Eindhoven University of Technology Library.

ISBN: 978-90-386-5306-8

NUR: 926

Keywords: Visual Perception, Chromatic Flicker, Temporal Contrast Sensitivity, Temporally Uniform Color Space, Dynamic Colored Light

Cover design: 隐身少女Alice

Printed by: ProefschriftMaken || www.proefschriftmaken.nl

Modeling the Temporal Behavior of Human Color Vision for Lighting Applications

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven,
op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens,
voor een commissie aangewezen door het College voor Promoties,
in het openbaar te verdedigen op maandag 5 juli 2021 om 13:30 uur

door

Xiangzhen Kong

geboren te Shandong, China

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

Voorzitter:	prof. dr. ir. A.W.M. Meijers
1 ^e promotor:	prof. dr. I. Heynderickx
copromotoren:	dr. I.M.L.C. Vogels dr. D. Sekulovski (Signify)
leden:	prof. dr. M. Fairchild (Rochester Institute of Technology) prof. dr. ir. W.L. IJzerman prof. dr. A.M.L. Kappers
adviseur:	prof. dr. P. Hanselaer (KU Leuven)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

DEDICATED

To

My parents

KONG Fanhe and SUN Jianli

Thank you for giving me endless love, unconditional support, and invaluable educational opportunities. I am deeply indebted.

My sister

KONG Chen

Thank you for being such a great sister who has always been a caring playmate and always takes care of me and understands me.

My grandparents

WEI Chunlan (*nainai*), KONG Qingrui (*yeye*, late), SUN Guanxin (*laoye*, late), WEI Qingying (*laoniang*, late)

Thank you all for taking care of me since I was born. Dear *nainai*, thank you for making so many brand-new shoes and warm winter clothes when I was a kid and thank you for always taking very good care of me. The delicious homemade *baozi* was an important part of my childhood memory. Dear *yeye*, I will always miss you. You were such a humorous and righteous person. You were my teacher who taught me calligraphy, although I am terrible at it. But I still remember the fragrance of the ink when it spreads over the old newspaper. You were a great playing buddy who drove me everywhere in town with my moving castle - your blue pedicab. You were my guardian in my life. I wish I could be your grandson again if there were an afterlife. Dear *laoye*, you were the first one in our big family who went abroad. Thank you for always encouraging us to grab the educational opportunities and take them seriously. Thank you for planting a seed in my heart that I wanted to go to university since I was a little kid. Dear *laoniang*, thank you for always telling me not to waste any food and respect the hard work of the farmers. You have always been a role model in this. I always leave the dining table with an empty plate.

Jeroen Drummen

Thank you for introducing me to real Dutch life: I have enjoyed cycling when following the *fietsknooppunten*. I have enjoyed all the mashed potato dishes (well, well, some of them). I have enjoyed exploring the territory of Dutch food: *brood met kaas en appelstroop*, *brood met chocolade hagelslag*, brood met the world. Thank you for encouraging me to be more Dutch (I don't know what that really means). Thank you for being the special one, and thank you for making me loved by my Dutch family: Magda, Hendrik, and oma. Kusjes!

Contents

1	Introduction	1
1.1	From Conventional Lighting to LED Lighting	1
1.1.1	White Light and Colored Light	2
1.1.2	Static Light and Dynamic Light	2
1.2	High-quality Dynamic Light	3
1.2.1	Color Path	3
1.2.2	Smoothness	3
1.2.3	Controllable Speed	4
1.3	General Research Goal	4
2	Color Perception	5
2.1	Human Visual System and Color Vision	5
2.1.1	Trichromacy	6
2.1.2	Color Opponency	6
2.1.3	Magnocellular, Parvocellular, and Koniocellular Pathways	8
2.2	Color Spaces	10
2.2.1	Introduction	10
2.2.2	Overview of Commonly Used Color Spaces	11
2.3	Temporal Color Changes	17
2.4	Research Questions and Structure of the Thesis	18
3	Perceived Speed of Changing Color in Chroma and Hue Directions in CIELAB	21
3.1	Introduction	22
3.2	Methods	23
3.2.1	Experimental Setup	23
3.2.2	Stimuli	24
3.2.3	Experimental Conditions	26
3.2.4	Participants	26
3.2.5	Procedure	26
3.2.6	Verification Measurements	28
3.3	Results	29
3.3.1	Calculation of PSEs	29
3.3.2	Comparison Between 10 PSEs	31
3.3.3	Questionnaire Analysis	32
3.4	Comparison with Available Color Models	33
3.4.1	Calculation of Slopes in CIELAB	33
3.4.2	Calculation of Slopes in LMS Cone Fundamentals	35

3.5	Discussion	36
3.5.1	Optimization of Color Spaces	36
3.5.2	Limitations and Future Work	37
3.6	Conclusions	39
4	Assessing the Temporal Uniformity of CIELAB Hue Angle	41
4.1	Introduction	42
4.2	Methods	43
4.2.1	Apparatus	43
4.2.2	Adaptation Condition and Stimuli	44
4.2.3	Experimental Conditions	46
4.2.4	Observers	47
4.2.5	Procedure	47
4.3	Results	48
4.3.1	Possible Bias	48
4.3.2	Calculation of PSEs and Confidence Intervals	49
4.4	Discussion	50
4.5	Conclusions	53
5	An Experimental Comparison of Threshold Methods for Chromatic Flicker De- tection: Accuracy, Precision and Efficiency	55
5.1	Introduction	56
5.2	Psychophysical Methods	57
5.2.1	Classical psychophysical methods	57
5.2.2	Adaptive Psychophysical Methods	59
5.2.3	Comparison of Threshold Methods	61
5.2.4	Objectives and Hypotheses	61
5.3	Methods	63
5.3.1	Experimental Design	63
5.3.2	Experimental Setup	63
5.3.3	Stimuli	63
5.3.4	Methodology	64
5.3.5	Participants	67
5.3.6	Procedure	67
5.3.7	Analyses	69
5.4	Results	71
5.4.1	Results of Method 1 – Classical Staircase with Yes-no Task	71
5.4.2	Results of Method 2 – Weighted Staircase with 2AFC Task	72
5.4.3	Results of Method 3 – Method of Constant Stimuli with 2AFC Task	73
5.4.4	Results of Method 4 – Method of Adjustment without Reference	74
5.4.5	Results of Method 5 – Method of Adjustment with Reference	75
5.4.6	Comparison of Methods	77
5.4.7	Questionnaire Analysis	80
5.5	Discussion	81
5.5.1	Effect of Base Color and Frequency	81
5.5.2	Comparison of Methods	82
5.5.3	Possible Limitations	84

5.6	Conclusions	85
6	Modeling Contrast Sensitivity for Chromatic Temporal Modulations	87
6.1	Introduction	88
6.2	Method	88
6.2.1	Experimental Setup	89
6.2.2	Stimuli	89
6.2.3	Participants	90
6.2.4	Procedure	91
6.3	Analyses and Results	93
6.3.1	Modeling of TCSFs	93
6.4	Discussion	96
6.5	Conclusion	97
7	Modeling Sensitivity to Chromatic Temporal Modulations Using Individual Cone Fundamentals	99
7.1	Introduction	100
7.1.1	Background - the Need for a Temporally Uniform Color Space	100
7.1.2	Individual Differences	101
7.1.3	Isoluminance	102
7.1.4	Contrast Sensitivity for Chromatic Temporal Modulations	102
7.1.5	Research Objectives	103
7.2	Methods	104
7.2.1	Apparatus	104
7.2.2	Generating the Stimuli	104
7.2.3	General Procedure	106
7.3	Experiment 1: Isoluminance	107
7.3.1	Participants	107
7.3.2	Stimuli	107
7.3.3	Procedure	108
7.3.4	Results	108
7.4	Experiment 2: Sensitivity to Chromatic Temporal Modulations	112
7.4.1	Participants	112
7.4.2	Stimuli	113
7.4.3	Procedure	113
7.4.4	Results	114
7.5	Discussion	121
7.5.1	Estimating Individual Cone Fundamentals	122
7.5.2	Quantitative Evaluation of Temporal Color Spaces in Terms of Local and Global Uniformity	123
7.5.3	Frequency-consistency	124
7.5.4	Individual Differences	125
7.5.5	Adaptation	125
7.6	Conclusion	125

8 General Discussion	127
8.1 Limitations and General Implications	129
8.1.1 Choice of Color Space	129
8.1.2 Adaptation	130
8.1.3 Individual Differences and the Use of Individual Isoluminance	130
8.1.4 Field of View, the Interaction of Central Vision and Peripheral Vision. . .	131
8.1.5 Limitations of the Hardware	132
8.1.6 Psychophysical Experiments	132
8.2 Future Work	132
8.2.1 Improved Method for Estimating Individual Cone Fundamentals	133
8.2.2 Improved Color Space	133
8.2.3 Adaptation	133
8.3 Summary of Contributions of the Current Thesis and General Implications . .	134
Summary	149
List of Publications	153
Curriculum Vitæ	155
Acknowledgements	157



Introduction

1.1. FROM CONVENTIONAL LIGHTING TO LED LIGHTING

Light plays an important role in our life. Apart from stimulating the visual system, which enables us to see the world around us, light also affects our emotions and biological clock, and as such, it is beneficial to human health, well-being, and productivity ([142, 118, 7]). When indoor spaces cannot sufficiently be illuminated by natural daylight, artificial light sources are used. In the past, several lighting technologies have been developed, resulting in, for example, incandescent light and fluorescent light. More recently, LED (light-emitting diode) lighting is being used both for general lighting and decorative lighting.

Conventional illumination technologies such as incandescent lighting convert the electric current to heat. When the filament is heated up to a specific temperature, it glows and produces visible light, making it very energy-inefficient, as only less than 5% of the energy is converted into visible light. So far, choices for the filament were rather limited (i.e., tungsten, halogen), and as a result, the spectral distribution of the emitted light was restricted. Moreover, these light sources typically respond relatively slowly to the input current. LEDs, on the contrary, offer a number of advantages compared to these conventional lighting technologies. They typically have a long lifetime, a fast temporal response, a low power consumption and they are produced from environmentally friendly materials. Furthermore, they have a wide controllability in terms of spectral power distribution, spatial distribution, color temperature, and temporal modulation. As a consequence, they can produce highly saturated colors at improved spatial and temporal resolution. Specifically, this characteristic has enabled the design of smart lighting that is adjustable to specific environments and requirements ([118]) and has facilitated the design of radically new, dynamic and attractive lighting atmospheres.

To understand the impact of dynamic and colored lighting on people, the concept of Human Centric Lighting (HCL) has gained attention. Dynamic colored LED lighting can, for example, be optimized to improve people's cognitive performance, which has been proven in educational settings such as classrooms [95]). Furthermore, LED-based HCL has a richer variation of possibilities to create emotionally appealing and stimulating atmospheres, leading to improved positive emotions ([102]).

1.1.1. WHITE LIGHT AND COLORED LIGHT

People usually prefer white light over colored light as the general illumination of an indoor space, especially when reading, writing, or chatting ([146]). Obviously, white light comes closest to natural daylight, which was the only available light source (apart from fire) for many centuries. Later incandescent lamps and fluorescent lamps have been the standard light sources for more than one and a half centuries, making us used to whitish light for our daily-life activities. White light, in this respect, is considered as a collective concept: it can differ both in illuminance level and in color temperature. Since the work of Kruithof ([65]) on the preference of general lighting, extensive research has been done to explore the influence of the intensity and color temperature of light on people's task performance and affective state. For example, Boyce et al. [8] found that elderly people perform better under light with a higher CCT (6500 K) than a lower CCT (3000 K) when examining the charts of Landolt rings. Yang and Jeon [162] found that a CCT of 4000 K leads to better working memory than a CCT of 3000 K for male students. In the study of Huang et al. [46], it was found that a CCT of 4300 K resulted in significantly better focused and sustained attention than 2700 K and 6500 K. Compared to research on white light, research on the effects of colored light on human beings is still limited. This might be due to the fact that it is unnatural and less preferable to perform daily tasks under colored light. However, colored light, when applied as decorative lighting, can contribute significantly to the atmosphere of an environment ([146]) and, as such, also influence people's mood ([43, 68]).

1.1.2. STATIC LIGHT AND DYNAMIC LIGHT

Both white light and colored light can be static or dynamic. In the past, general lighting has been mainly static, and therefore it has received a lot of attention in literature. Dynamic light is the application of dynamics in the intensity, chromaticity and/or spatial distribution of the light. LEDs have enabled inexpensive ways to create dynamic lighting effects, and as a consequence, understanding the perception and preference of these dynamics is drawing more and more attention from both academia and industry.

For academia, LED-based systems have provided researchers more well-controlled stimuli to explore the human visual system and the perception of dynamic light. Before the era of LED lighting, the creation of dynamic light required complicated mechanical devices such as a combination of mirrors and an episcotister ([140]) or a spinning polarizer ([55]) to generate flickering light stimuli. Nowadays, temporally modulated light is easily generated with LED-based systems. For example, Sekulovski et al. [121] used LEDs driven with pulse width modulation (PWM) to generate stimuli for temporal color perception. Eisen-Enosh et al. [19] adopted a 5000 Hz sampling rate device to drive an LED for generating stimuli to measure the critical-fusion frequency. Murdoch [96] adopted a multi-primary LED system to research human visual adaptation to temporally dynamic light. Those research opportunities would be almost impossible without an LED-based system.

For industry, dynamic colored light has been used in several applications, for instance, to enhance the alertness of office workers ([7]), to create appealing light atmospheres ([40, 79, 97, 145]), and to enhance the immersive experience of displays, such as in the Philips Ambilight TV ([115]). The Ambilight TV creates dynamic light effects around the television screen that are aligned with the video content on the screen. It is an example of the enhanced functionality that dynamic colored light can bring to displays and media presentations, aimed at providing a more immersive and attractive viewing experience to

consumers ([120]).

1.2. HIGH-QUALITY DYNAMIC LIGHT

Dynamic light should have the following properties in order to be attractive to observers: (1) the light should show pleasant color combinations, (2) the light should change smoothly over time in luminance and/or chromaticity, and (3) the light should change at the desired perceived rate of change ([97, 121, 120]). Recent research recommends that the lighting design should also take into consideration the non-image-forming (NIF) effect of light on our circadian rhythms ([56]). Although important, we consider the latter aspect outside the scope of the current thesis.

1.2.1. COLOR PATH

Changes in luminance are one-dimensional: i.e., from low to high intensity or vice versa. Changes in chromaticity, on the other hand, describe a path in a two-dimensional chromaticity space. For example, the chromaticity can change along a straight line between two colors or along a curved line. The specific path used to go from one color to another is one of the factors that determine the attractiveness of dynamic colored light. Vogels, Sekulovski, and Rijs [144] found that the appealingness of chromaticity changes along a straight path between two colors strongly depends on the type of color space that was used to define the path, such as RGB and CIELAB. Hartog [40] performed an experiment-based study to measure participants' preferences for dynamic light and found that color and speed are the most important parameters that people want to change to create dynamic light settings. Despite the importance and still limited knowledge on this topic, we decided not to focus on the preferred color path in this thesis, but rather to focus on the perceived speed of the color transition.

1.2.2. SMOOTHNESS

LEDs are usually digitally controlled, which means that changes in luminance and/or chromaticity are implemented as small steps. A disadvantage of this property is that it may introduce visible artifacts, such as jerkiness. Abrupt changes in light are common in specific scenarios such as concert lighting and theatre lighting, but for most applications, they are not experienced as pleasant. Therefore, light should preferably change smoothly over time from one condition to another, both in luminance and in chromaticity. The preferred speed of the transition may strongly depend on the context, but for any given speed, the light is considered of high quality only if the transition is implemented without perceived flicker or abrupt changes. Sekulovski et al. [121] found that the visibility threshold of smoothness, i.e., the maximum color difference between two successive colors that is allowed in order to perceive a temporal color transition as smooth, is about ten times smaller for lightness changes than for chroma or hue changes in CIELAB. This means that CIELAB, which is a commonly used color space for spatial color perception, is not a useful space to predict the perception of dynamic colored light. Today, no color spaces are available that accurately predict the visibility of color differences over time.

1.2.3. CONTROLLABLE SPEED

Another property that influences the attractiveness of dynamic light is the perceived speed of the color transition, i.e., the magnitude of the change in luminance and/or chromaticity per unit of time ([120]). Preferably, this speed is controllable from slow transitions (e.g., over hours) to fast transitions (e.g., within seconds). Currently, the perceived speed of color transitions has not been investigated systematically.

1.3. GENERAL RESEARCH GOAL

In this dissertation, we are mainly interested in the perceived speed of dynamic colored light. This is a challenging topic since the perception of temporal color differences is far less understood than the perception of spatial color differences. In the past, multiple color spaces have been designed to define perceptual equality for spatial color differences. However, a specifically designed color space for temporal color differences does not exist yet. Since spatial properties of a visual stimulus are processed in different areas of the visual cortex than temporal properties [57], it is unlikely that metrics developed for spatial color perception are accurate in predicting temporal color changes. Thus, we need a human-vision based color space to describe the perception of temporal changes in colored light. To approach this challenge, we first give in **Chapter 2** a more detailed overview on human color perception and the characteristics of existing spatial color spaces.

Color Perception

2.1. HUMAN VISUAL SYSTEM AND COLOR VISION

The general structure of the human visual system has been known for decades, and that information is available in many standard textbooks (see, for example, [53, 148, 88]). Color vision starts with the absorption of light in the retinal cone photoreceptors, which are long-wavelength sensitive (L-cones), middle-wavelength sensitive (M-cones), and short-wavelength sensitive (S-cones). The cones are much more concentrated in the central yellow spot (known as the macula) of the retina: L- and M-cones are concentrated in the center of the retina, with a density that decreases towards the periphery (Figure 2.1), while S-cones peak in density just outside the center fovea [11]. Rods are absent in the *fovea centralis* (i.e., a 0.3-mm diameter rod-free area in the center of the macula), and rise to a high density away from it, spreading over a large area of the retina. Rods are mainly active at low light levels (i.e., between 10^{-6} and $10^{-3.5}$ cd/m²). Vision dominated by rods is known as *scotopic vision*, while vision dominated by the cones at higher luminance levels (i.e., 10 to 10^8 cd/m²) is known as *photopic vision*. At intermediate luminance levels (i.e., $10^{-3.5}$ to 10 cd/m²), vision is facilitated by both rods and cones, which is referred to as *mesopic vision*.

Cones and rods transduce electromagnetic energy into electrical voltages, which are then transformed into action potentials by a complicated network of cells in the retina [33]. The network of cells consists of horizontal cells, bipolar cells, amacrine cells, and ganglion cells, as shown in Figure 2.2 for the cones only [21].

The signals from the retina are sent to the lateral geniculate nucleus (LGN) and then to the occipital cortex via axons, which are known as optic radiation, the geniculocalcarine tract, the geniculostriate pathway, or posterior thalamic radiation. This sketched high-level outline from light in the eye to visual information in the brain may look straightforward, but actual color perception is complicated, and not all details of the process are understood yet. This is already true for a static image, let alone the complexity of understanding the visual perception of a temporally modulated colored light signal. Concepts that are nonetheless relevant for the perception of (dynamic) colored light are: trichromacy, color opponency, and parvocellular, magnocellular, and koniocellular pathways. These concepts are explained in more detail below.

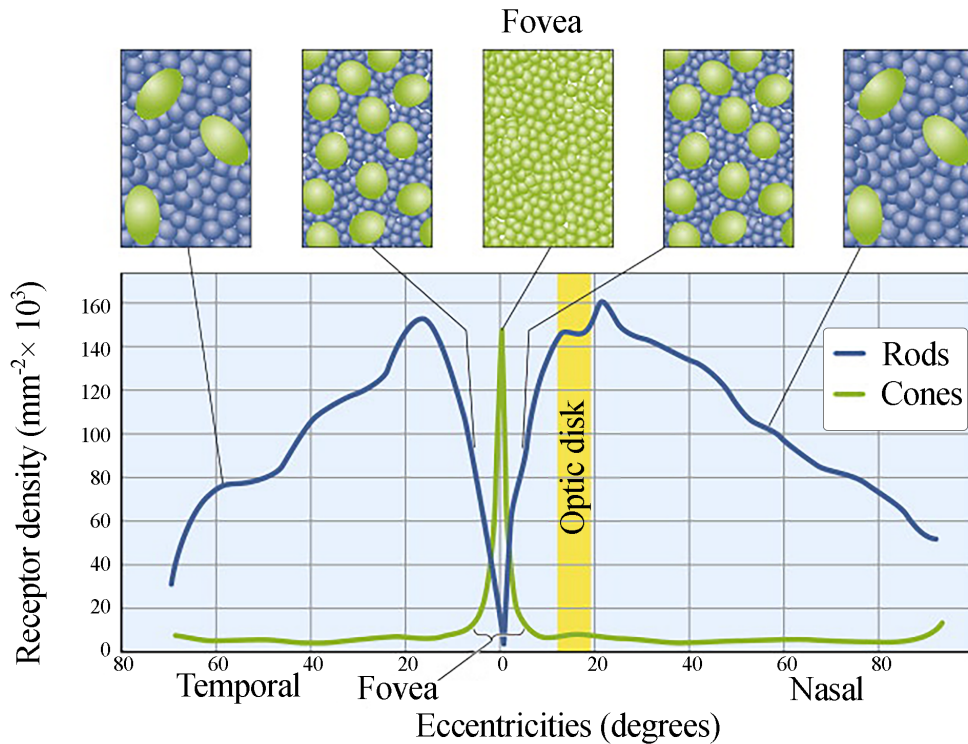


Figure 2.1: Distribution of rods and cones in the human retina (illustrated in a modified version from [98]). The five boxes at the top show a cross-section through the outer segment of the photoreceptors at different eccentricities. Cones have an increased density and decreased diameter of their outer segment towards the center of the retina (i.e., the fovea).

2.1.1. TRICHROMACY

The trichromacy theory was proposed by Thomas Young in 1802 and further developed by Herman von Helmholtz to finally become known as the Young-Helmholtz trichromatic theory. The main reason that human vision is trichromatic is because of the existence of three types of cone photoreceptors that together are responsible for color perception. These photoreceptors have a univariant output, which means that any photon that is absorbed by the photoreceptors has the same effect on its firing rate as any other absorbed photon, independent of its wavelength. What does differ between the three types of cones, however, is the probability that a photon with a given wavelength is absorbed. The relation between this probability and the wavelength of the photon is described in the so-called spectral sensitivity function. Figure 2.3 shows the spectral sensitivity function for the L-cones, M-cones, and S-cones separately. A consequence of the univariant nature of cone photoreceptors is that the photoreceptor response is determined both by the intensity of the light (which is related to the number of photons) and by the wavelength of the photons (which determines the chance of a photon being absorbed).

2.1.2. COLOR OPPONENCY

The Young-Helmholtz trichromatic theory is able to explain and describe various fundamental properties of the human visual system. For instance, two light fields with a different spectral distribution can be perceived as having the same color, which is called metamerism. Color matching experiments have shown that a test field composed of a single wavelength

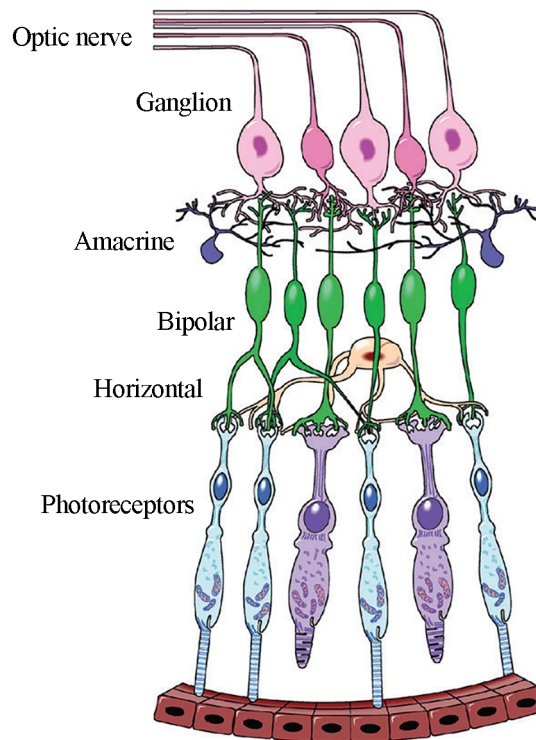


Figure 2.2: A diagram that shows the vision signal path from the receptors to the optical nerve. Adapted from [21].

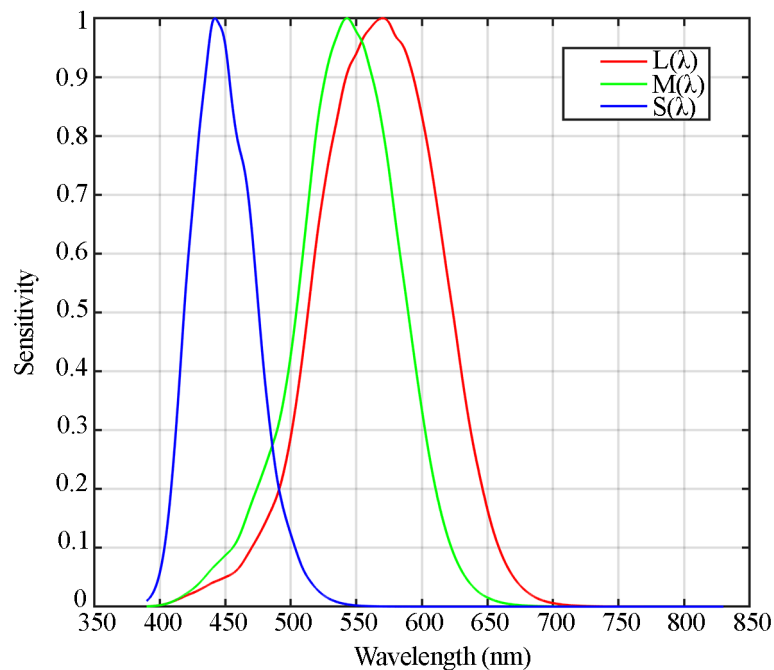


Figure 2.3: Normalized cone spectral sensitivities for the S-cones (blue), M-cones (green) and L cones (red) (Dataset retrieved from: <http://www.cvrl.org/>).

can be matched with an adjacent field by adjusting the relative intensity of three light sources

of different wavelengths. On the other hand, the trichromatic theory cannot explain several other human vision phenomena, such as the afterimage effect. For instance, after looking at a red light for several minutes, a white object appears to be greenish. In 1874 Ewald Hering proposed the color-opponent theory to explain this effect. According to this color-opponent theory, color is encoded as three relative cone outputs, as shown in Figure 2.4. The red-green channel is given by the L-M+S contribution, while the yellow-blue channel is given by the S-(L+M) contribution. L+M+S represents the luminance channel (or the achromatic channel). Since, in many cases, the S-cone input is negligible for the red-green and achromatic channels, simply L-M and L+M are accepted to represent those two channels. However, some researchers argue that also S-cones contribute to the luminance channel under certain conditions [109]. In the most simple model of the color-opponency theory, all cone contributions are evenly weighted. However, the cone weights for the different channels are not fixed and vary, among other things, between foveal and peripheral vision [114]) and among individuals [130].

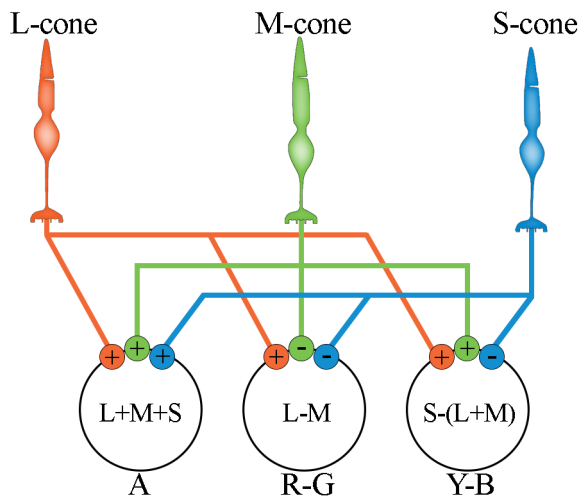


Figure 2.4: Photoreceptor processing and color opponency. Schematic illustration of the encoding of cone signals into opponent-color signals. A stands for the achromatic channel, R-G stands for the red-green channel, while Y-B stands for the yellow-blue channel.

At the time that Ewald Hering proposed his color-opponency theory, it conflicted with the Young–Helmholtz trichromatic theory. Later we discovered that both theories could be combined since they describe different stadia of visual processing. Because of their complementary value, both color representations (i.e., LMS and the three opponent channels) currently co-exist. One should realize that although there has been a lot of psychophysical evidence that supports the color-opponent theory (for example, the existence of separable ON and OFF processing pathways [91]), the physiological basis of the color-opponent mechanisms remain unclear [124, 161]. In addition, for explaining adaptation effects, the color-opponency theory is not entirely satisfactory [91].

2.1.3. MAGNOCELLULAR, PARVOCELLULAR, AND KONIOCELLULAR PATHWAYS

In order to know how temporal information is processed in the visual system, we need to describe the properties of the visual pathways. Recently, anatomical and electrophysiological studies have revealed that there are three pathways from the retina to the LGN with different spatial and temporal properties: the magnocellular (MC), parvocellular (PC), and koniocel-

lular (KC) pathway [73, 99, 128]. MC, PC, and KC pathways are formed of morphologically distinct cellular layers that receive information from different types of retinal ganglion cells and project to different layers in the primary visual cortex [86]. Although evidence has accumulated that simple and direct associations between physiological responses and perceptual properties are difficult to establish [86], there has been a growing consensus about the link between the parvocellular pathway and the chromatic L-M channel and the link between the magnocellular pathway and the luminance channel (L+M). The koniocellular pathway is sensitive to S-cones but is less understood yet.

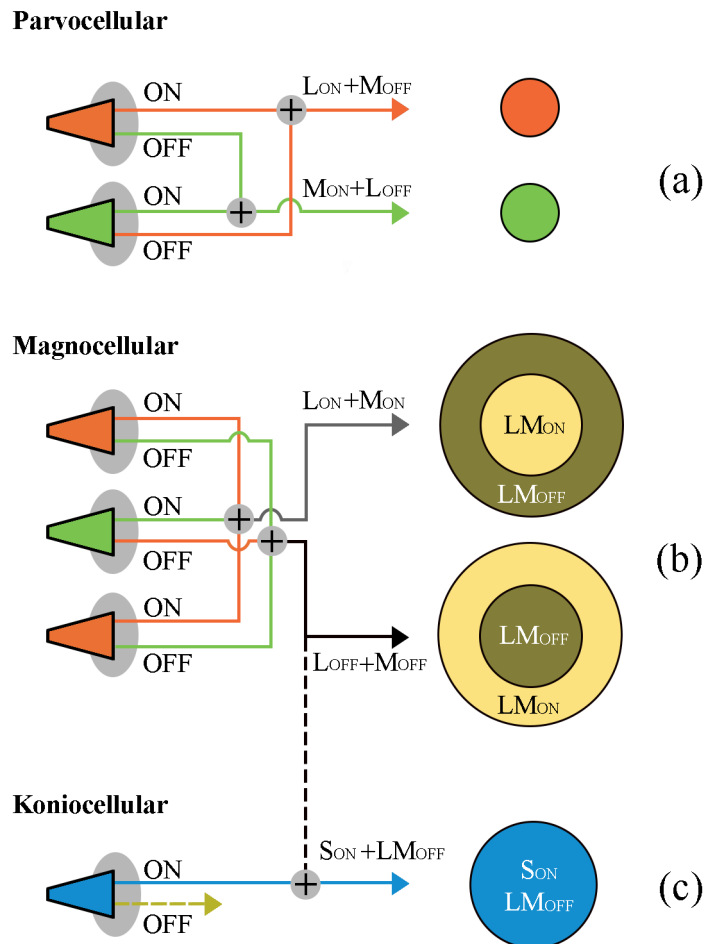


Figure 2.5: A simplified model of parvocellular (PC), magnocellular (MC), and koniocellular (KC) pathways. (a) The PC pathway responds to the relative, opposing variations in L- and M-cone input (i.e., L - M input, which roughly corresponds to “red-green” variations); (b) The MC pathway responds to the overall change in L- and M-cone input (i.e., L + M input, which roughly corresponds to luminance variations); (c) The KC pathway responds to the relative, opposing variations in the S-cone input against the overall L- and M-cone input (i.e., S - (L + M) input, which roughly corresponds to “yellow-blue” variations). The sizes of the receptive fields at the output of the pathways are approximate and not to scale. Adapted from [131].

The magnocellular pathway has a high temporal resolution, low spatial resolution, high contrast sensitivity, and no color sensitivity, whereas the parvocellular pathway has a moderate temporal resolution, low contrast sensitivity, excellent spatial resolution, and high red-green color sensitivity [127]. The general network model of the pathways is described in [131, 32] and shown in Figure 2.5.

2.2. COLOR SPACES

2.2.1. INTRODUCTION

2

To enable precise color reproduction in both scientific studies and real-world applications, there is a clear need to describe the perception of color in a mathematically sound way, which is exactly what the field of colorimetry is doing [21]. Before the introduction of colorimetry, color order systems such as the Munsell system provided a means to specify colors in terms of their appearance in values as lightness, hue, and chroma (i.e., Munsell Value, Munsell Hue, and Munsell Chroma). However, the lack of a sound mathematical framework in color order systems limited the possibilities of reproducing the same color perception under different viewing conditions.

Based on colorimetry, color scientists developed multiple color spaces, i.e., mathematical models that aim to describe perceived color in a quantitative way. First of all, these models assume a standard observer, which represents the average behavior of a specific target group based on empirical data. Secondly, an important requirement of these color spaces is that the Euclidian distance between two points in a color space is a representative measure for the perceived color difference between these points, independent of the location and direction in the color space. In order to measure and express perceived color difference, the concept of just-noticeable difference (JND) is used. The JND for a specific color corresponds to the smallest deviation from this color that can be distinguished by a human observer. In an ideal color space, the JND is equal everywhere in the color space and, therefore, does not depend on the location in the color space. A color space that fulfills this property is called perceptually uniform.

Unfortunately, the earlier color spaces (such as CIE 1931 XYZ and CIE 1976 Uniform Chromaticity Scales (UCS), see also Section 2.2.2) were not perceptually uniform, and so Euclidian distances in one part of the space resulted in a larger or smaller perceived color difference than in another part of the space. Furthermore, as Fairchild [21] pointed out, these color spaces could not provide us with a numerical description of the (relative) appearance of colors or how the appearance of colors changed under different viewing conditions. Therefore, in the late 1970s, color appearance models were developed. CIELAB and CIELUV were the first color spaces that could be considered as a color appearance model because they took adaptation to a reference white into account, and they expressed color appearance in terms of hue, saturation (or chroma), and lightness. Although the latter might sound simple, in practice, this turns out to be more difficult because these color characteristics are not perceptually independent. For example, Stalmeier and Weert [129] noted that when the saturation of an object increases, the perceived brightness also increases, which is known as the Helmholtz-Kohlraush effect. Conversely, the perceived saturation of a color increases with its luminance level, also known as the Hunt effect [25]. Hence, there is not yet a perfect model for describing color appearance in terms of three perceptually independent values.

Colorimetry has long focused on perceptual differences between colors in the spatial domain, i.e., for color patches presented next to each other, and so in spatial coexistence. The resulting color spaces can be considered as *spatial* color spaces and can be used, for example, to make smooth spatial color transitions. In this thesis, we focus on dynamic colors, in other words, on temporal color transitions. Hence, to describe the perception of temporal color changes, we need a *temporal* color space. There is, however, currently no validated color space for predicting the perception of color changes over time.

In the next section, we give a summary of the development of commonly used spatial

color spaces, since they can be used as a source of inspiration for the development of a temporal color space.

2.2.2. OVERVIEW OF COMMONLY USED COLOR SPACES

Several color spaces have been defined in the past, each with their own advantages and disadvantages. Below we give a brief historical overview of the development of color spaces.

CIE 1931 XYZ COLOR SPACE

In the 1920s, two color matching experiments were independently carried out by Wright [160] and Guild [37]. The color matching experiments were carried out using a 2-degree bipartite field, where one side of the field was illuminated by monochromatic light with a known wavelength, while the other side was illuminated with a combination of three lights using additive color mixing, each representing a primary source of red, green and blue light. The task of the participant was to adjust the relative amount of the three primary light sources to match the monochromatic light (as shown in Figure 2.6). The results of these color matching experiments described how to weight the three primary colors in order to match any monochromatic color, resulting in functions that are now known as Color Matching Functions (CMFs).

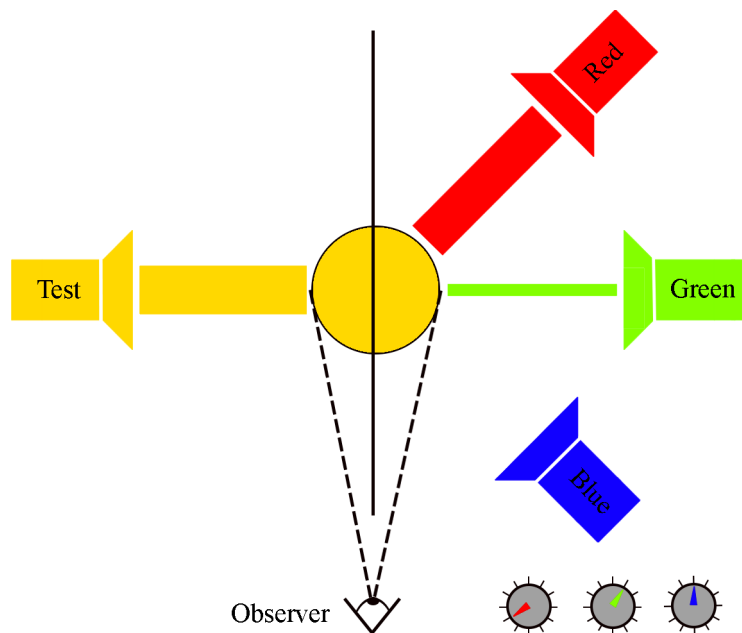


Figure 2.6: The classical color matching experiment: a process of determining a unique RGB triplet for each stimulus. At the left side of the bipartite field, a uniform stimulus with monochromatic light of a known wavelength is projected. At the right side, a perceptually equal uniform stimulus needs to be generated by additive mixing of three lights with known primaries.

Although different primaries and different intensity levels were used in Wright and Guild's experiments, their results were summarized as the standardized CIE RGB color matching functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$ by the CIE in 1931. The mathematical calculations were based on Grassman's first law of color mixture [35], which states that any color can be matched by a linear combination of three other colors, provided that none of those three can be matched by a combination of the other two. Specifically, the CMFs are the amounts of three

monochromatic primaries at standardized wavelengths of 700 nm (red), 546.1 nm (green), and 435.8 nm (blue) needed to match the monochromatic test primary. These CMFs form the basis of the CIE RGB color space (Figure 2.7).

2

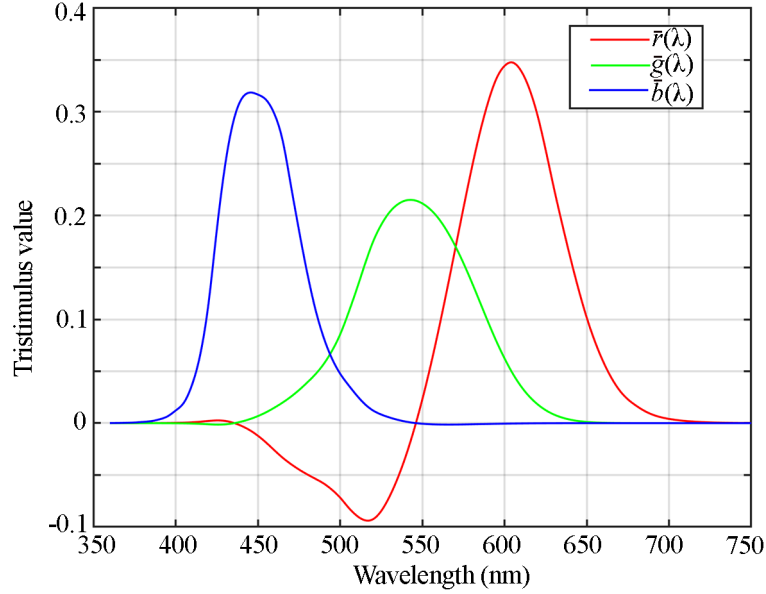


Figure 2.7: The color matching functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$ of the CIE 1931 RGB color space.

A disadvantage of the CIE 1931 RGB space, like any other RGB space, is that it depends on the specific choice of the primaries, making it device-dependent. As a result, the same RGB triplet may represent a different color in a different device. In addition, the negative values in $\bar{r}(\lambda)$ were inconvenient in computations. Therefore, three imaginary primaries X , Y , and Z were defined such that they eliminated the negative values in the CMFs, i.e., $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ (Figure 2.8). X is chosen to be a mix of the three CIE RGB color matching functions, such that they are non-negative. Z is quasi-equal to $\bar{b}(\lambda)$ of the CIE RGB color matching functions, while Y is the luminance when calculated with the photopic luminous efficiency function $V(\lambda)$ of the standard observer. The advantage of setting Y as the luminance is that for any given Y -value, the XZ -plane contains all possible chromaticities at this luminance. The transformation from RGB to XYZ is linear and given by Equation 2.1.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49000 & 0.31000 & 0.20000 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00000 & 0.01000 & 0.99000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.1)$$

In the XYZ color space, each color is defined as a weighted combination of the three imaginary primaries X , Y , and Z , where the weights are labeled as the *tristimulus values*. Since a 2-degree bipartite field was used, this set of CMFs is also known as the 1931 CIE 2° CMFs. By normalizing the XYZ values by the sum of X , Y , and Z , the CIE xyY color space was derived, using the following equations:

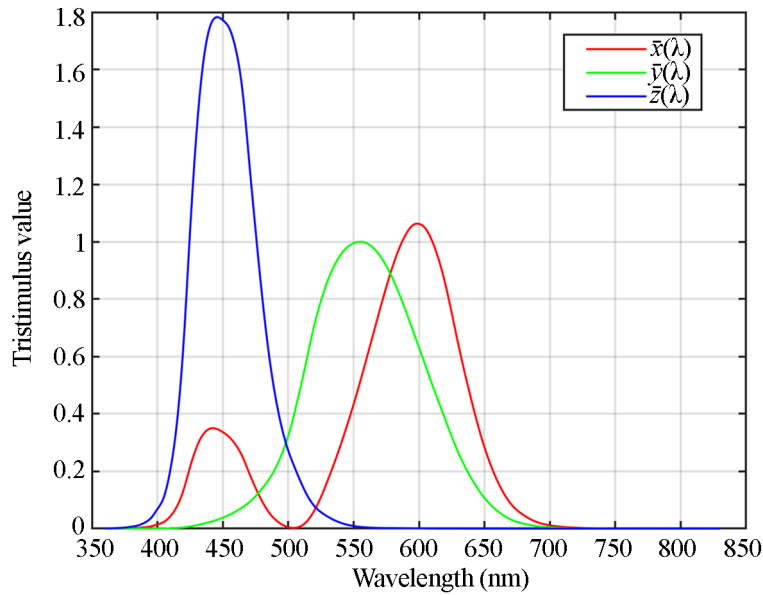


Figure 2.8: CIE 1931 XYZ color matching functions, which consist of $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$.

$$\begin{cases} x = \frac{X}{X + Y + Z} \\ y = \frac{Y}{X + Y + Z} \\ z = \frac{Z}{X + Y + Z} = 1 - x - y \end{cases} \quad (2.2)$$

In the notation xyY , Y is the luminosity, or luminance of a color, while the two parameters x and y are defined in a plane perpendicular to this luminance axis and denote the chromaticity of a color. This xy -plane is known as the CIE 1931 xy chromaticity diagram, shown in Figure 2.9.

The CIE 1931 XYZ is a more useful color space to describe color in absolute terms than any of the RGB color spaces since the primaries are fixed and, therefore, independent of the properties of the device that produces the color. However, the JNDs in color are not the same across the entire color space. MacAdam [84] demonstrated this non-uniformity by fitting ellipses through the color points that were perceived as identical with respect to a given base color (represented by the center of the ellipse) at different locations in the xy chromaticity diagram (Figure 2.10(a)). These ellipses are now known as the MacAdam ellipses. Many color scientists have formulated adaptations to the CIE 1931 XYZ color space since then, aiming at making the MacAdam ellipses more circular.

CIE 1976 UCS COLOR SPACE

One attempt to make the CIE 1931 XYZ color space more perceptually uniform for colors at approximately the same luminance is the CIE 1976 UCS color space (see Figure 2.10(b)). The luminance corresponds to Y , while the chromaticity is expressed in u' and v' and can be calculated from the tristimulus values of XYZ (or from the chromaticity coordinates xy) according to the following equations:

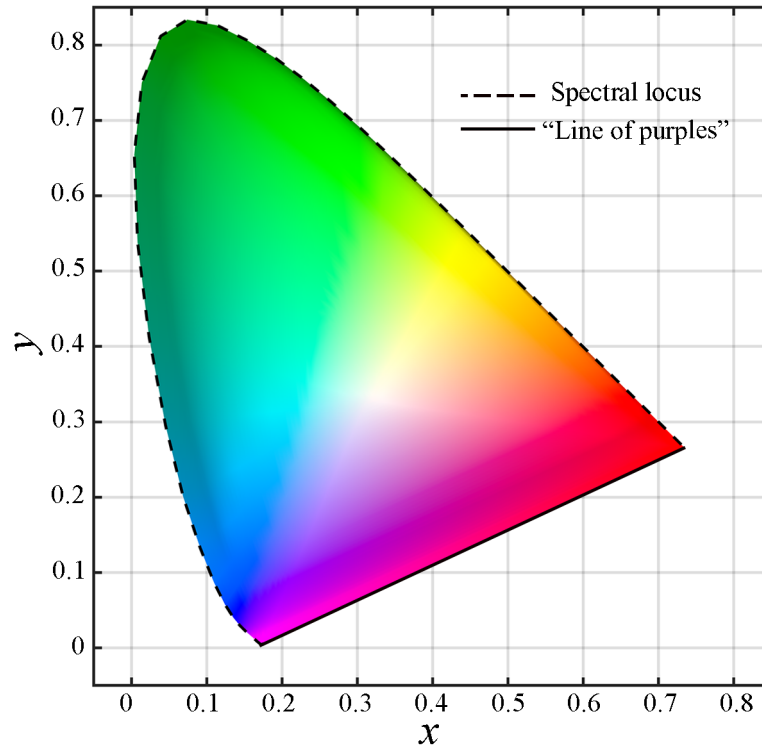


Figure 2.9: CIE 1931 xy chromaticity diagram. The edge of the diagram, indicated by the dashed line, is called the spectral locus and represents monochromatic light. The solid line represents "line of purples". The colors on this line and the colors inside the diagram can only be produced by color mixture.

$$\begin{cases} u' = \frac{4X}{X + 15Y + 3Z} = \frac{4x}{-2x + 12y + 3} \\ v' = \frac{9Y}{X + 15Y + 3Z} = \frac{9y}{-2x + 12y + 3} \end{cases} \quad (2.3)$$

CIELAB COLOR SPACE AND CIELCH COLOR SPACE

Another attempt to create a color space that is based on the quantification of perceived color differences resulted in the CIE 1976 $L^* a^* b^*$ color space, referred to as CIELAB. An important distinction with earlier spaces is that it is a relative color space, where colors are described in relation to their viewing conditions. As such, CIELAB can be considered as a color appearance model. CIELAB uses lightness (L^*) to describe the brightness of a color relative to the brightness of a similarly illuminated white area, called the white point. The chromaticity is represented by a^* and b^* , where a^* refers to the redness or greenness, and b^* refers to the yellowness or blueness of the color. It is important to realize that the a^* and b^* axes do not exactly correspond to the aforementioned red-green and blue-yellow

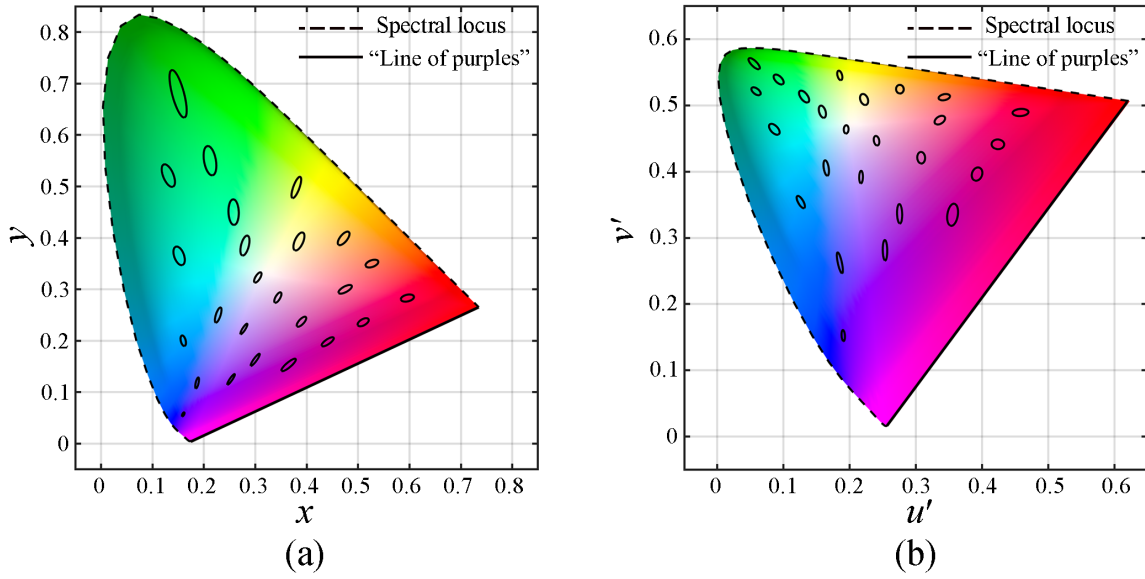


Figure 2.10: (a) MacAdam ellipses in CIE 1931 xy chromaticity diagram (b) MacAdam ellipses in CIE 1976 UCS chromaticity diagram. All the chromaticity points at any ellipse are 5 JNDs different from the center point.

opponent channels. The conversion from CIE 1931 XYZ to CIELAB is as follows:

$$\begin{cases} L^* = 116f\left(\frac{Y}{Y_r}\right) - 16 \\ a^* = 500\left(f\left(\frac{X}{X_r}\right) - f\left(\frac{Y}{Y_r}\right)\right) \\ b^* = 200\left(f\left(\frac{Y}{Y_r}\right) - f\left(\frac{Z}{Z_r}\right)\right) \end{cases} \quad (2.4)$$

where

$$f(t) = \begin{cases} \sqrt[3]{t} & \text{if } t > 0.00856 \\ 7.78707t + 0.13793 & \text{otherwise} \end{cases}$$

and (X_r, Y_r, Z_r) refers to the tristimulus values of the white point.

CIELAB provides a relatively good computational model for predicting perceived (spatial) color differences (see Figure 2.11). The color difference metric is commonly abbreviated as ΔE_{ab}^* ("Delta E") and is generally implemented as the Euclidean distance in CIELAB, defined as:

$$\Delta E_{ab}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \quad (2.5)$$

Since perceived colors are usually not described in terms of redness or blueness but rather in terms of hue (or the tint of a color), saturation (or the degree of chromaticity of a color), and its brightness (or lightness), the CIE $(L^* C^* h)$ color space, referred to as CIELCH, was defined as a derivative of the CIELAB color space. In this color space, colors are characterized by their lightness (L^*), chroma (C^*), and hue (h), which are defined as:

$$\begin{cases} L^* = L^* \\ C^* = \sqrt{a^{*2} + b^{*2}} \\ h = \begin{cases} \arctan\left(\frac{b^*}{a^*}\right) & \text{if } \arctan\left(\frac{b^*}{a^*}\right) > 0 \\ \arctan\left(\frac{b^*}{a^*}\right) + 360^\circ & \text{otherwise} \end{cases} \end{cases} \quad (2.6)$$

The CIE has also standardized several color difference metrics based on the CIELCH color space, namely the CIE94 and CIEDE2000 color difference metrics. The CIE94 ΔE_{94}^* introduced different weights for the difference in L^* , C^* , and h between two colors. The CIEDE2000 ΔE_{00}^* added an additional rotation term to account for the non-uniformities in the bluish region of the color space, as visible in Figure 2.11.

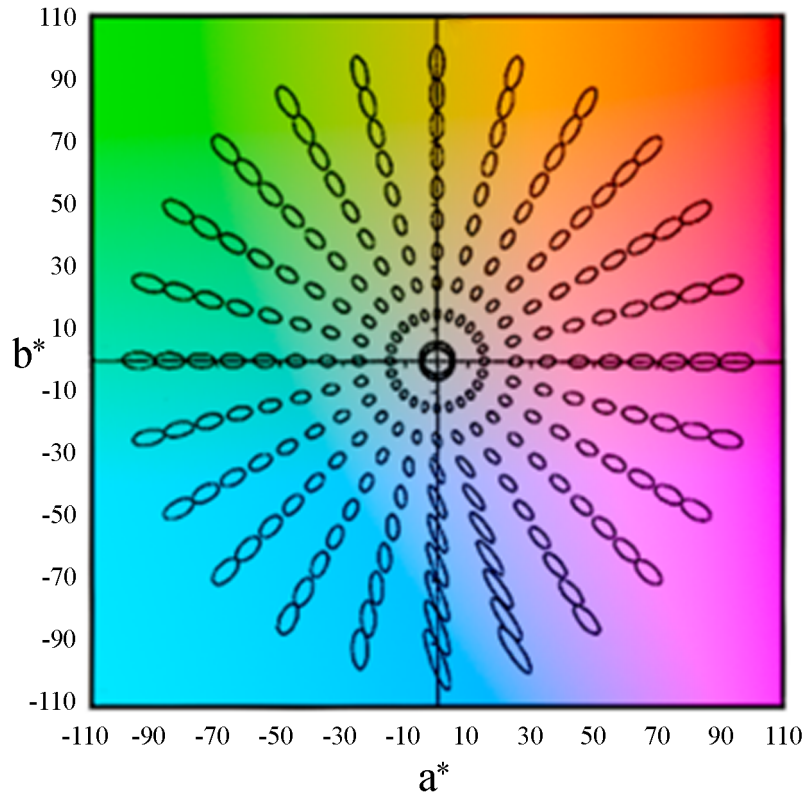


Figure 2.11: MacAdam ellipses in CIELAB (Figure retrieved and adapted from https://fujiwaratko.sakura.ne.jp/infosci/lab_e.html).

PHYSIOLOGICALLY BASED COLOR SPACE

The above-mentioned color spaces were all derived from psychophysical experimental data. As such, they do not reflect the underlying physiological mechanism of how a light signal is processed. In an attempt to remain closer to the color-opponent theory, Derrington, Krauskopf, and Lennie [12] studied how S-, M-, and L-cones contributed to the opponent responses in the LGN of monkeys. They defined a color space with three orthogonal axes, of which the intersection corresponded to the white point. The axes represented the luminance

axis ($L + M$), the S-cone excitation axis [$S - (L + M)$], and the L- or M-cone excitation axis ($L - M$), which are illustrated in Figure 2.12. The resulting color space is now commonly dubbed as the DKL space. It has proven its value in the vision science community, especially for its easier manipulation of stimuli. For example, modulation along the $L - M$ axis is invisible to S-cones Rinner and Gegenfurtner [108]. Similarly, modulation along the $S - (L + M)$ cone axis changes the excitation of S-cones only and is invisible to L- and M-cones.

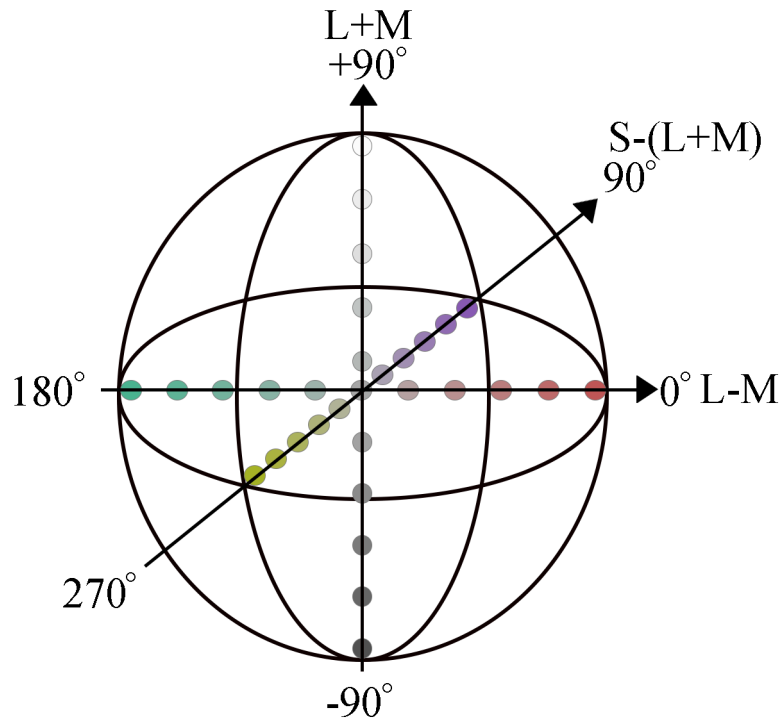


Figure 2.12: The axes of the DKL color space, namely $L-M$, $S-(L+M)$, and $L+M$.

Similar to the CIE 1931 xy chromaticity diagram, one can also define a physiologically based chromaticity diagram. In the MacLeod-Boynton chromaticity diagram [85], the two axes are $l (= L/(L+M))$ and $s (= S/(L+M))$, as shown in Figure 2.13.

2.3. TEMPORAL COLOR CHANGES

As mentioned in **Chapter 1**, artificial light can be made dynamic by changing its intensity, chromaticity, or spatial distribution. These temporal changes can be periodic or non-periodic and can have various speeds. For example, temporal changes can be very *fast*, such that only the colors at the start and the end of the transition are perceived, and not the intermediate colors. At the other end of the scale, temporal changes can be very slow, such that the transition is hardly noticeable, as might happen for changes in daylight. In this thesis, we are mostly interested in transitions with intermediate speeds.

Existing literature mostly focuses on the two extreme rates of change. There is an extensive body of literature on the visibility (and annoyance) of very *fast* periodic changes, also known as flicker. More technically, the term “flicker” is defined by the CIE as “the impression of unsteadiness of visual perception induced by a light stimulus whose luminance or spectral distribution fluctuates with time” [48]. Both luminance flicker and chromatic flicker have been extensively studied. Most literature concerns luminance flicker, mainly because

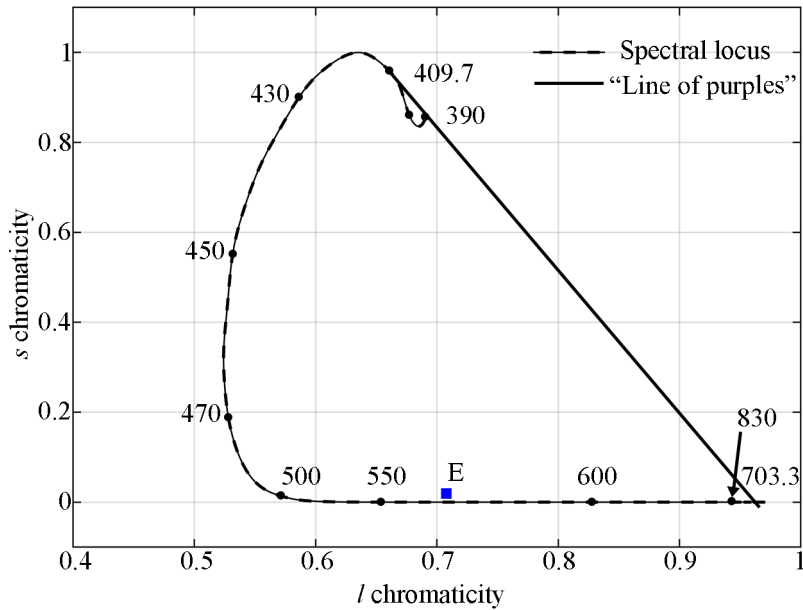


Figure 2.13: MacLeod-Boynton chromaticity diagram. The values along the spectral locus are in nanometer. The line of purples is defined as the straight-line tangent to the spectrum locus at a wavelength of 409.7 nm and ending at the spectrum locus with a wavelength of 703.3 nm.

human vision is far more sensitive to temporal changes in the achromatic channel than in the chromatic channels at high frequencies [136]. A useful way to visualize the sensitivity to fast periodic changes is the temporal contrast sensitivity function (TCSF), where the sensitivity to flicker (i.e., the inverse of the Michelson contrast of the visibility threshold) is a function of the frequency of the fluctuation. The sensitivity to luminance flicker is typically a bandpass-shaped function of temporal frequency, while the sensitivity to chromatic flicker exhibits a low-pass-shaped function.

For *intermediate* and *slow* transitions, smoothness is a more important characteristic [97, 120]. Sekulovski et al. [121] introduced “Delta-E-ab per second” to describe the speed of temporal color transitions. In order to be a useful metric, any temporal color transition with discrete color steps of $1 \Delta E_{ab}^*$ per time unit would be perceived as smooth and as having a constant rate of color change during the entire transition. Sekulovski et al. [121], however, demonstrated that the maximum $\Delta E_{ab}^*/s$ that is allowed to make a luminance transition perceptually smooth is a factor 10 smaller than for transitions along a chromaticity direction. This implies that ΔE_{ab}^* is not a uniform measure for describing temporal color differences.

Sekulovski et al. [121] also investigated the relationship between the perception of slow and fast transitions. They measured the smoothness thresholds for linear temporal light transitions and the visibility thresholds of chromatic flicker at the same base color. They found a high correlation between both types of dynamic transition, suggesting that knowledge on flicker perception can be used to develop a model predicting the perception of gradual changes in light.

2.4. RESEARCH QUESTIONS AND STRUCTURE OF THE THESIS

Different chapters are based on different publications and are not arranged in a chronological order.

We argued above that the latest attempts to define a perceptual uniform color space (such as CIELAB) were designed to predict the appearance of colors when stimuli are presented side by side. We have addressed the concept of *uniformity* many times, and it has become clear that the perceived difference between two colors should be proportional to the distance (usually Euclidian) within a perceptually uniform color space. Sekulovski et al. [121] have shown that CIELAB (or its derivative CIELCH) is not uniform for describing the perception of temporal color differences. Hence, there is a need for a new temporal color space. To the best of our knowledge, this challenge is not yet addressed in the literature. Therefore, the long-term goal of this project is to develop a temporal uniform color space, in order to create smooth dynamic light effects with well-defined perceived speed. This challenge has two aspects: (1) each color transition should be perceived as smooth, and (2) the perceived speed of the transitions should correspond to the intended speed, which may be constant throughout the whole transition for some applications, but it may also change during the transition for other applications. In this thesis, we provide data that will contribute to the formulation of a uniform temporal color space in order to fulfill both requirements of dynamic light. The outline of this thesis is as follows.

In **Chapter 3** and **Chapter 4** we describe two studies that measured the perceived speed of several temporal colored light transitions with a fixed change in ΔE_{ab}^* per second. We expected that transitions with the same ΔE_{ab}^* per second are not necessarily perceived as having the same speed, as a similar effect was found for smoothness perception and flicker [121]. In the experiments, we generated linear colored light transitions in CIELAB around several base colors that changed either along the hue direction (i.e., constant lightness and chroma) or along the chroma direction (i.e., constant lightness and hue). The experiments addressed our **first research question**, being: what is the effect of direction (i.e., chroma and hue) and location in the color space (i.e., base color) on the perceived speed of a colored light transition? Once a perceptually uniform temporal color space is available, it should be able to predict the outcome of these experiments.

In **Chapter 5**, we started to collect data for developing a temporal color space. Even though our focus was on transitions with intermediate speed, we decided to measure the sensitivity to *fast* transitions, more specifically to chromatic flicker. First of all, the effects of base color and modulation direction were found to be similar for both relatively *slow* and *fast* colored light transitions despite the difference in absolute threshold [121]. Secondly, participants found the task to determine the visibility threshold for chromatic flicker much easier than for smoothness. Chromatic flicker thresholds are expressed as the smallest amplitude of a sinusoidal modulation that results in perceived flicker. Ideally, this amplitude is the same for a modulation in any direction around any base color in the color space, which means that the requirement of uniformity should be met both locally and globally. This implies that a large amount of data should be collected, in order to assure that those data are representative for the whole color space. In order to avoid a high burden for the participants of our experiments, we needed to establish an efficient but accurate measurement method. Although several psychophysical methods have already been compared in literature, the conclusions were not consistent and depended on the type of stimulus. Therefore, we compared several commonly used psychophysical methods to quantify chromatic flicker thresholds. This resulted in our **second research question**, being: what is an efficient yet accurate method for measuring large amounts of subjective data, and in our particular case for measuring perceived flicker of colored light modulations?

In **Chapter 6**, the most optimal measurement method was used to collect data on the visibility of chromatic flicker. As there are countless possibilities to create chromatic flicker, e.g., by varying the sinusoidal modulation in multiple directions at multiple temporal frequencies around multiple base colors, we needed to find a logical and smart way of limiting the experimental parameters. Previous research has shown that the effect of frequency on flicker sensitivity can be modeled as a temporal contrast sensitivity function (TCSF), which generally corresponds to an exponential function. When the TCSF is known for a number of base colors that are equally distributed in the color space and a number of modulation directions, we can select a few representative frequencies in follow-up studies and reduce the number of stimuli. Therefore, our **third research question** was: can we use one and the same model to describe the effect of temporal frequency on the visibility of chromatic flicker for a wide range of chromatic modulations, and, do the parameters of the model depend on base color and modulation direction? To address this research question, we measured the visibility threshold of chromatic flicker for nine base colors at four modulation directions (orthogonal in CIE 1976 UCS chromaticity diagram) and seven temporal frequencies for three participants.

The data could be modeled by several exponential functions with a high goodness-of-fit, where the two parameters of the model (i.e., slope and intercept) depended on base color and modulation direction. This means that theoretically, only two temporal frequencies are needed for measuring the TCSF for a specific combination of base color and modulation direction. It also shows that the measure used to quantify contrast sensitivity is not appropriate to define one TCFS for the entire color space. Therefore, our **fourth research question** was: what is a suitable contrast measure to unify the parameters of the TCSFs? In addition, large individual differences were found, not only in the absolute thresholds but also in the rate (or slope) at which the thresholds decreased with frequency. This led to our **fifth research question**: can we reduce individual differences in TCSF by taking into account individual cone spectral sensitivities? In **Chapter 7**, we addressed both questions by collecting more chromatic flicker thresholds for more participants. The data were analyzed by taking into account estimated individual cone spectral sensitivities and by comparing different contrast measures.

The above-mentioned studies had provided us new and useful insight on constructing a temporal uniform color space. In **Chapter 8**, we describe a general framework and requirements for constructing such a temporally uniform color space. In addition, we reflect on our research questions, summarize the general conclusions, and discuss ideas for further research.

3

Perceived Speed of Changing Color in Chroma and Hue Directions in CIELAB

In dynamic LED lighting, the perceived speed of changing color is an important concept; however, there exists no suitable temporal color space. In a psychophysical experiment, we compared the perceived speed of periodic temporal transitions in CIELAB chroma and hue directions around five base colors [the five Munsell hues: 5R (red), 5Y (yellow), 5G (green), 5B (blue), and 5P (purple)]. The experiment was conducted in a light laboratory, with the main illumination stimulus subtending a visual angle of 101×77 deg. In sequential paired presentations, observers were asked to identify which transition appeared faster, and points of subjective equality between transitions were computed. The speed of transitions was defined in CIELAB $\Delta E_{ab}^* / s$ was shown to be temporally non-uniform; uniformity was improved using a modified color space based on speeds in the DKL space of Derrington et al. [12].

This chapter is copied with (slight) adaptations from Journal of the Optical Society of America A, Volume 36, Number 6, 1022-1032 (2019) as: "Perceived Speed of Changing Color in Chroma and Hue Directions in CIELAB", Xiangzhen Kong, Michael J. Murdoch, Ingrid Vogels, Dragan Sekulovski, and Ingrid Heynderickx [61].

3.1. INTRODUCTION

The development of light-emitting diode (LED) technology has enabled easy and inexpensive ways to create dynamic colored light. As a result, dynamic colored light is nowadays used in several applications, for instance, to enhance the alertness of office workers [7], to create appealing light atmospheres [145, 40, 97, 79], and to enhance the immersive experience of displays, such as in the Philips Ambilight TV [115]. In most of these applications, the light has to fulfill (at least) the following two requirements in order to be attractive to observers: the light should change smoothly over time, in luminance and/or chromaticity, and the light should change at the desired rate of change (or speed) [97, 121, 120]. Several studies have investigated how people perceive smoothness of dynamic light generated with limited temporal resolution [121], subtlety of dynamic light [97], and its attractiveness [145]. However, it is not known yet how to describe the perceived rate of a temporal change in color. For instance, if we want to generate a temporal light transition with a constant perceived rate of change, we need to know what is perceived as an equal amount of change in color per time unit.

Color science has long focused on perceptual differences between colors in the spatial domain, leading to metrics such as CIEDE94 and CIEDE2000 that describe the magnitude of a color difference for spatially separated colors. These metrics are formulated in the perceptually uniform CIELAB color space. Although the uniformity of the CIELAB color space is not perfect (since it depends, among other things, on hue [67]), this color space is considered useful for describing how to make smooth spatial color gradients. Since there is currently no validated model for temporal color perception available, it is most reasonable to use CIELAB to describe temporal color changes as well, and to express the speed of a color transition in terms of $\Delta E_{ab}^* / s$. This was done for the first time in 2007 by Sekulovski et al. [121], who measured the maximum color difference between successive colors at which a temporal color transition was still perceived as smooth. When the smoothness threshold was expressed in terms of $\Delta E_{ab}^* / s$, a linear relationship was found with the temporal frequency. However, the smoothness threshold was a factor of 10 lower for temporal changes in lightness compared to changes in hue and chroma. This means that the CIELAB color space is quite non-uniform for describing temporal color differences. Therefore, the question remains how to describe the speed of a temporal color transition in a perceptually uniform way.

Existing literature on speed perception cannot be used to answer this question since it manipulates other (for the purpose of our study irrelevant) stimulus characteristics. For example, a few studies have measured the sensitivity to changing luminance for achromatic Ganzfeld stimuli (e.g., [58]). Here, log luminance per min was used as a metric for describing the temporal change in luminance. However, in this study, we are interested in the sensitivity to changing chromaticity. In addition, a large amount of literature exists on motion perception for spatiotemporal stimuli, for which speed can be expressed as degrees per second (where “degrees” refer to the spatial distance in terms of visual angle). However, we want to know the perceived speed of a light transition for a (ideally) homogeneous light stimulus. Since the spatial structure of the stimulus is constant, there is no change in visual angle.

Our long-term goal is to develop a universal metric for describing the perceived rate of a temporal color change. In this study, we aim to present new data on speed perception using temporal stimuli defined in CIELAB and evaluate how well existing color models, both CIELAB and cone spaces including LMS and Derrington, Krauskopf, and Lennie (DKL) [12]

describe the results. Therefore, we conducted a psychophysical experiment in which we measured the perceived speed of temporal linear color transitions in chroma and hue at constant lightness in the CIELAB color space. In order to quantify the perceived magnitude of the speed, several stimuli (with different speeds) were compared to a reference stimulus (with a fixed speed), and the point of subjective equality (PSE) in speed was determined. As expected, none of the models just mentioned were suitable for predicting perceived rate of color change. We propose a straightforward adjustment to these models and evaluate which model is the most promising for further improvement and validation.

3.2. METHODS

The experiment was designed to find the PSE of speed between different color transitions. To do so, a two-interval forced-choice task was used. More particularly, participants were instructed to compare the perceived speed of two temporal color transitions (i.e., a reference stimulus and a comparison stimulus). The transitions varied either in CIELAB chroma or hue direction around a base color point. Five base color points were used, corresponding to the principal Munsell hues 5R (red), 5Y (yellow), 5G (green), 5B (blue), and 5P (purple) [21]. We limited ourselves to two types of stimulus pairs: (1) one stimulus was modulated in the chroma direction and the other stimulus in the hue direction around the same base color point, referred to as *CH*-comparison; (2) the two stimuli were both modulated in the hue direction around two different base color points, referred to as *HH*-comparison. A third possible type of comparison, namely *CC*-comparison (i.e., two stimuli both modulated in the chroma direction around two different base color points), were omitted in this experimental design to limit the experimental time for the subjects, and because higher-amplitude *CC* stimuli would have resulted in out-of-gamut colors in some conditions. The method of constant stimuli was used to find the PSE of speed with respect to the reference stimulus by fitting a psychometric function through the percentages “the comparison stimulus is faster than the reference stimulus”, averaged over the participants. The experiment used a full-factorial within-subject design with 26 participants.

3.2.1. EXPERIMENTAL SETUP

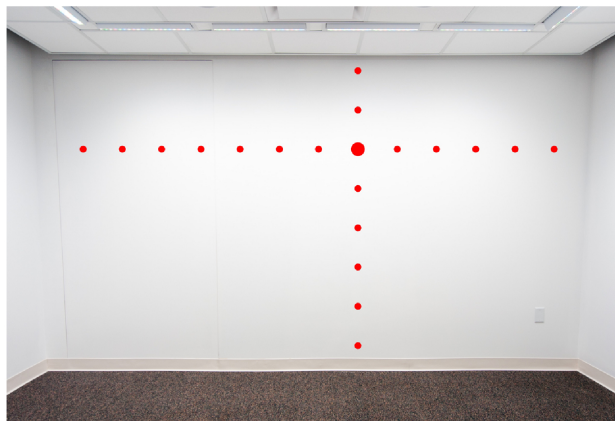


Figure 3.1: Photograph of the DVA Lab with a neutral white 4000 K light setting. The red dots added to the photo indicate the positions on the wall that were measured; the larger dot shows the location of the characterization measurements.

The experiment was conducted in a laboratory (see Figure 3.1) that is designed to study dynamic color perception and adaptation, dubbed the Dynamic Visual Adaptation (DVA) Lab of the Munsell Color Science Lab of Rochester Institute of Technology. The DVA Lab is a 3.66 m × 4.27 m × 2.44 m (length × depth × height) room with a variegated gray carpet and four white walls without a window. All the walls are painted with a matte white paint having an average reflectance of 93%, and the floor has an average reflectance of 7%. One of the long walls consists of three sliding wooden panels 2.4 m wide, which can be opened. The room is equipped with a five-primary (red, green, blue, mint green, white: *RGBMW*) LED lighting system that can be addressed at 40 Hz to create smooth temporal changes. The lighting system consists of 14 Philips SkyRibbon wall-washing fixtures that are mounted in the ceiling to illuminate the lab's walls, resulting in a smoothly non-uniform illumination pattern. The uniformity was assessed via spectral measurements at positions spaced 30 cm apart, noted by red dots in Figure 3.1; the larger red dot is the location of the brightest region of the wall, where the spectral characterization data used to model the lighting system were measured. Vertically, the point above and the two points below the reference were all above 70% of the reference luminance, and the top and bottom points were 38% and 40% of the reference. Horizontally, the central nine points (the middle 2.8 m of the wall) all measured above 86% of the reference luminance, and the outermost points reached 38%. Chromaticity uniformity is described in Section 3.2.6 below, after the color stimuli are explained. More information about the lab is described elsewhere [96]. We intentionally chose a full room setting instead of the often used 2° or 10° stimuli, since dynamics in general lighting are more realistically represented with a larger field of view; similarly, we allowed participants to look freely around the illuminated wall.

3.2.2. STIMULI

The stimuli were periodic temporal color transitions around a base color point in the chroma or hue direction of the CIELAB LCh color space [see Figure 3.2(a)]. The color changed every Δt seconds with a step size S over a full period of 3 s, which was found in pilot experiments to be both long enough to be visibly changing and short enough to allow many stimulus presentations in a manageable experiment length. The color transition was shown as five full periods of a triangular wave to allow for easier judgements on the speed of the transition. The sharp edges of the triangular wave were smoothed, as indicated by the regions R^* in Figure 3.2(b). This was done in order to avoid abrupt changes that might influence the perception of speed, as suggested in [121].

Each cycle had a fixed number of discrete steps ($N = 120$), and each step was shown for a fixed period $\Delta t = 0.025$ seconds. Thus, the stimulus lasted for 15 (i.e., $5 \times 120 \times 0.025$) seconds. 10% (i.e., $p = 0.1$) of this stimulus duration was used for smoothing the abrupt color change. Hence, the stimulus consisted of $N \times (1 - p)$ number of steps with a fixed step size S and $N \times p$ steps with variable step size $S^* = \frac{S}{2^i} (1 \leq i \leq N \times p)$. The color transition speed (CTS) is defined as the speed of the linear part of the transition and is calculated as

$$CTS = \frac{S}{\Delta t} \approx \frac{A}{T} \quad (3.1)$$

with A being the amplitude of the transition and T the duration of one half of the cycle (i.e., 1.5 s). In our experimental design, the duration of the transition was fixed, and the speed of the color transition was implemented by changing the amplitude of the transition [see

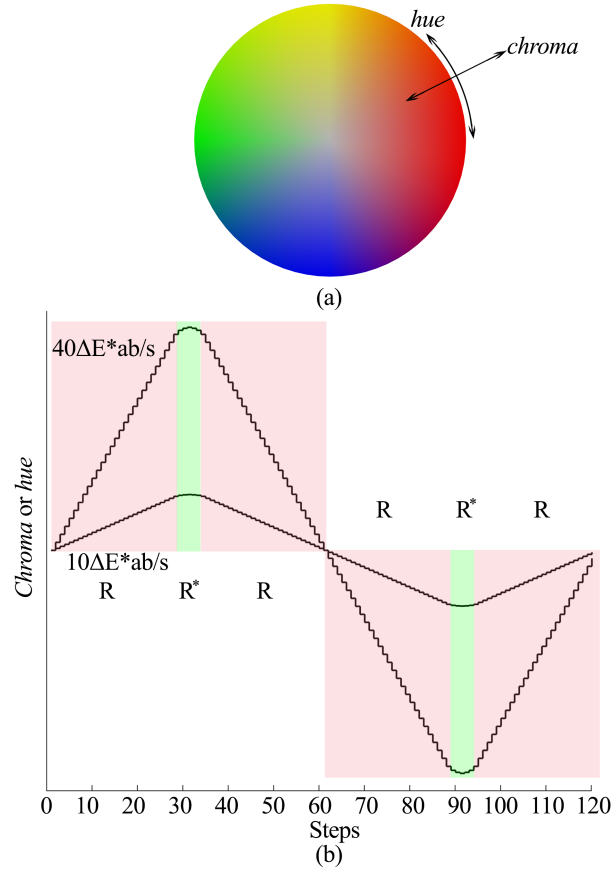


Figure 3.2: (a) Chroma and hue directions in the Ch plane of CIELAB LCh color space. (b) Temporal color transition as a function of time with a speed of $10 \Delta E_{ab}^*/s$ or $40 \Delta E_{ab}^*/s$. The regions marked *R* have a fixed step size, and hence a fixed speed. The region marked *R** is adapted to smooth the abrupt change in color.

Figure 3.2(b)]. The amplitude A can be calculated as follows:

$$A = S \times \left[N \times (1 - p) + N \times p \times \left(1 - \frac{1}{2^{N \times p}} \right) \right] \quad (3.2)$$

Each color transition varied around a base color point, which corresponded to a principal Munsell hue: 5R (red), 5Y (yellow), 5G (green), 5B (blue), or 5P (purple). The color points of the principal Munsell hues were calculated as follows: first, 10° XYZ tristimulus values were computed from the reflectance spectra [42] of the matte Munsell color chips with the principal hues under Illuminant C. Because Munsell colors are defined to be uniform under Illuminant C, Von-Kries chromatic adaptation was used to find the corresponding color points under 4000 K, which was the white point of the neutral light setting used in the experiment. These XYZ tristimulus values were transformed to CIELAB LCh, with the adaptation white of 4000 K at a wall luminance level of 200 cd/m^2 , and the average CIELAB h for each group of Munsell hues was found. Each of the five base color points was defined at its average CIELAB hue with a chroma C^* of 60 and a lightness L^* of 90. The corresponding CIELAB hue angles are 27 for 5R, 81 for 5Y, 159 for 5G, 221 for 5B, and 302 for 5P. Each color transition varied around the base color point at a constant speed in either the chroma or hue direction. Five color transition speeds were used: 2.5, 5, 10, 20, and $40 \Delta E_{ab}^*/s$. Since the transitions were described in the CIELAB LCh color space, first the corresponding speeds in

terms of chroma and hue were calculated, using

$$\Delta C^* = \Delta E_{ab}^* \quad (3.3)$$

$$\Delta H^* = \arccos\left(1 - \frac{\Delta E_{ab}^{*2}}{2 \times C^{*2}}\right) \quad (3.4)$$

3

Then the L^* , C^* and h values were transformed to XYZ values, and finally to the $RGBW$ values that were used to control the LED luminaires. The XYZ to $RGBW$ transformation model that was used, with a white mixing ratio of 0.85, has been described in more detail in [96].

All the stimuli were within the gamut of the system. During the experiment, all the walls were illuminated; however, the seated participants were asked to look at the wall in front of them.

3.2.3. EXPERIMENTAL CONDITIONS

In each trial, two color transitions were shown sequentially, in random presentation order: a reference stimulus and a comparison stimulus. In CH -comparisons, the reference stimulus was modulated in the hue direction around one of the five base color points at a speed of $10 \Delta E_{ab}^*/s$, whereas the comparison stimulus varied around the same base color point in the chroma direction at one of the five speeds. Hence, there were 25 (i.e., 5 hues \times 5 speeds) CH -comparisons. In HH -comparisons, both stimuli were modulated in the hue direction. The reference stimulus varied around a hue of 5R (i.e., $h = 27^\circ$) at a speed of $10 \Delta E_{ab}^*/s$ the comparison stimulus varied around one of the five base color points at one of the five speeds. This means that there were also 25 (i.e., 5 hues \times 5 speeds) HH -comparisons.

3.2.4. PARTICIPANTS

Twenty-six adults (15 males and 11 females) volunteered to participate in the experiment. Their age ranged from 21 to 53 years ($M = 32.4$, $SD = 10.5$). All participants had normal color vision, as measured with the Ishihara test for color deficiency. Fourteen of them were wearing glasses, and six were wearing contact lenses. Most participants had more-than-average knowledge about color, while three of them had advanced knowledge about color. Ten participants did not have any experience with lighting experiments, while the others had different levels of experience.

3.2.5. PROCEDURE

The experiment was approved by the Institutional Review Board at RIT, the Human Subject Research Office. Before the start of the experiment, participants were asked to read and sign an informed consent form, in which they confirmed their voluntary participation.

When participants entered the experimental room, the light was set to a static white light with a correlated color temperature of 4000 K and a wall luminance of 200 cd/m². Participants were instructed to sit on a chair that was positioned at a distance of 1.5 m from the wall in front of them. The average eyesight level was about 1.5 m above the floor. No chin or forehead rest was used since we intended to have a more natural viewing setting. Then the experimenter read the instruction script that consisted of the following parts: welcome words, explanation of the necessity to do an Ishihara test, collection of

demographic information, explanation of the stimuli and task, illustration of the voice instructions, and a practice session. The participants were instructed that they were free to make head and eye movements, as long as they looked at the wall in front of them. During the practice session, each observer was asked to compare (1) the speeds of a hue change with $10 \Delta E_{ab}^*/s$ at 5R and a chroma change with $40 \Delta E_{ab}^*/s$ at 5R, and (2) the speeds of a hue change with $10 \Delta E_{ab}^*/s$ at 5R and a hue change with $2.5 \Delta E_{ab}^*/s$ at 5G. Each comparison was repeated twice. Six computer-generated voice instructions were used: **V1**, “First stimulus”; **V2**, “Second stimulus”; **V3**, “Which stimulus appears faster?”; **V4**, “Input recorded”; **V5**, “Invalid input, please input again”; **V6**, “Congratulations! You finished the experiment, thank you for your participation.”

The flow of presenting the stimulus pairs is shown in Figure 3.3, and consisted of voice-cued sequential stimuli interspersed with neutral 4000 K lighting that was shown for ten seconds. Research has shown that after ten seconds, chromatic adaptation is approximately 80% complete [108]. This means that participants started from approximately the same adaptation point. Participants were instructed to judge which of the two sequential color transitions appeared to be faster by pressing a button on a keyboard: “1” to indicate the first stimulus appeared faster and “2” to indicate the second stimulus appeared faster. If a participant happened to press a button that was neither “1” nor “2”, the voice instruction **V5** would be played until that participant pressed “1” or “2”. Then **V4** would be played and the program would proceed. Since participants had different levels of knowledge about color, the CIELAB LCh space was briefly explained, and it was pointed out that all the stimuli were only changing in either the chroma or hue direction at a constant luminance level. In addition, we emphasized that all stimuli had the same duration, and that a judgement regarding the speed of the transition could be made based on the slope of the transition. The concept of slope was explained by showing a similar figure to Figure 3.2(b). All the stimuli shown during the instruction and practice phase were part of the stimuli of the main experiment. Participants were allowed to ask questions during the instructions and they had the opportunity to repeat the practice session, if needed, before the start of the main experiment. During the practice session and the main experiment, participants wore a baseball cap to prevent them from looking directly into the luminaires mounted in the ceiling.

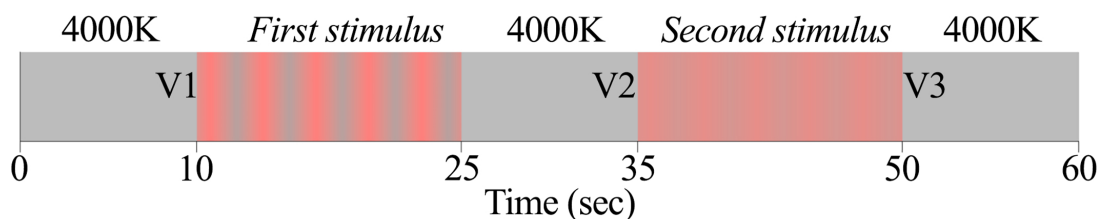


Figure 3.3: Illustration of the procedure for presenting one stimulus pair. **V1**, **V2**, and **V3** are the voice instructions.

The main experiment consisted of two sessions. In the first session, participants evaluated all 50 stimulus pairs [i.e., 25 *CH*-comparisons (5 hues \times 5 speeds) + 25 *HH*-comparisons (5 hues \times 5 speeds)]. In the second session, only the three intermediate speeds were presented, resulting in 30 comparisons [i.e., 15 *CH*-comparisons (5 hues \times 3 speeds) + 15 *HH*-comparisons (5 hues \times 3 speeds)]. As a result, the largest and smallest speeds of the

Table 3.1: Questionnaire after each session

Session	No.	Question
1	1	Do you think the task of the experiment is difficult? If so, what made it difficult?
	2	Where did you see the most visible color change? “Center”, “Periphery”, Or “No difference (between center and periphery)?”
	3	Did you notice anything that was not smooth (such as stepping, flickering, stuttering, etc.)? If so, please describe it (e.g., in which color).
	4	Do you feel the cap is influencing the judgement?
2	5	Do you feel the experiment is easier the second time than the first time?
	6	Any other comments?

comparison stimulus were presented only once, while the intermediate three speeds were repeated twice per participant. We decided to do this since the lowest and highest speed were visually quite different from the reference speed, and they were basically used (1) to test if participants understood the task by checking whether their responses were close to 100% correct, and (2) to make participants feel more confident about their performance by showing some easier trials. In each session, first all *CH*-comparisons were presented, blocked per base color, and then all *HH*-comparisons were shown. The presentation order of the base colors was based on a balanced Latin square design [125].

The second session was arranged at least half an hour after the first session. For some participants, the sessions took place on different days. All participants watched the demo stimuli and did the practice session again before the start of the second session in order to be sure that they had the same level of understanding. After each session, participants were asked to answer the questions shown in Table 3.1. They were encouraged to elaborate on their answers, and not simply respond with “yes” or “no”.

3.2.6. VERIFICATION MEASUREMENTS

All stimuli were verified (i.e., each step in the temporal color transition was measured) with a Photo Research PR-655 spectroradiometer. The location of the measurement was at 1.7 m above the floor, as indicated by the larger red circle in Figure 3.1. Among the 6000 (i.e., 50×120) measured steps, the average deviation from the corresponding intended colors was $2.34 \Delta E_{ab}^*$, with the largest deviation being $4.81 \Delta E_{ab}^*$ from the measurements, the actual transition speeds that were presented to the observers were further calculated, and the results showed that the speeds differed from the intended speeds by 2.17%.

The spatial chromaticity uniformity of the wall was measured at the same spatial locations shown in Figure 3.1, and Euclidean $\Delta u'v'$ values from the reference point were computed. Over all five base colors and the neutral white, at all spatial locations (total 6×19 measurements), the median $\Delta u'v'$ from the reference was 0.0010. The maximum difference was 0.0042 for the 5Y base color near the bottom of the wall.

3.3. RESULTS

3.3.1. CALCULATION OF PSES

First, we checked the reliability of the participant responses. In either *CH*- or *HH*-comparisons, the extreme speeds (i.e., 2.5 and 40 ΔE_{ab}^*) were visually quite different from the reference speed of 10 ΔE_{ab}^* . However, in 29 of the 520 cases [i.e., 26 participants \times 2 speeds \times 2 types of comparisons (*CH* and *HH*) \times 5 hues], participants either reported that the lowest speed was faster than the reference or the highest speed was slower than the reference. In particular, there was one participant with 6 out of 20 likely “incorrect” responses. Therefore, the data of this participant was removed for further analysis.

For each comparison type (*CH* and *HH*) and base color, the percentage of responses “the comparison stimulus is faster than the reference stimulus” was calculated over all participants and plotted as a function of the speed of the comparison stimulus. This resulted in 10 plots. For *CH* comparisons the reference was a transition of 10 ΔE_{ab}^* along the hue direction around the same hue as the comparison stimulus. For *HH* comparisons, the reference was a transition of 10 ΔE_{ab}^* along the hue direction around a (fixed) hue of 5R.

The resulting data were fitted with the psychometric function of Equation 3.5, using a generalized linear model (i.e., the `glmfit` function available from the Statistics Toolbox in MATLAB 2017b):

$$f = \frac{1}{1 + e^{-(a+b \times \log_{10} x)}} \times 100\% \quad (3.5)$$

In this equation, f is the percentage of responses, x corresponds to the speed, the parameter a defines the percentage when $x = 0$, and the parameter b is related to the slope of the psychometric function. The PSE is defined as the speed at which the percentage of responses “faster than the reference” equals 50%.

Figure 3.4 shows the ten psychometric curves fitted with the data of the 25 participants. In this figure, the speed is plotted on a logarithmic scale to improve the goodness of fit. The output of the logarithmic fit has a root mean square error of 1.68 $\Delta E_{ab}^*/s$, compared to the root mean square error of 2.97 $\Delta E_{ab}^*/s$ for the linear fit. We used a statistical technique based on Monte Carlo resampling [157] to estimate the variability in the parameters a and b of the psychometric function. To do so, we adopted a non-parametric bootstrap by treating all data points of all participants (for a given condition, i.e., *CH* or *HH* comparison and one base color) as if they reflected one underlying distribution. Next, 25 new data points were randomly chosen from this distribution for each of the five speeds and the PSE was calculated. This procedure was repeated 10,000 times, and from these data we calculated the average PSE and its confidence interval as the 2.5th and 97.5th percentile. These data are summarized in Figure 3.5.

The straightforward interpretation of the PSE is that, for example, a transition of 18.48 $\Delta E_{ab}^*/s$ along the chroma direction at 5Y is perceived as fast as a transition of 10 $\Delta E_{ab}^*/s$ along the hue direction at 5Y [Figure 3.5(a)]. At the same time, a transition of 6.12 $\Delta E_{ab}^*/s$ along the hue direction at 5Y is perceived as fast as a transition of 10 $\Delta E_{ab}^*/s$ along the hue direction at 5R [Figure 3.5(b)]. To describe the data in a more compact way, we use the

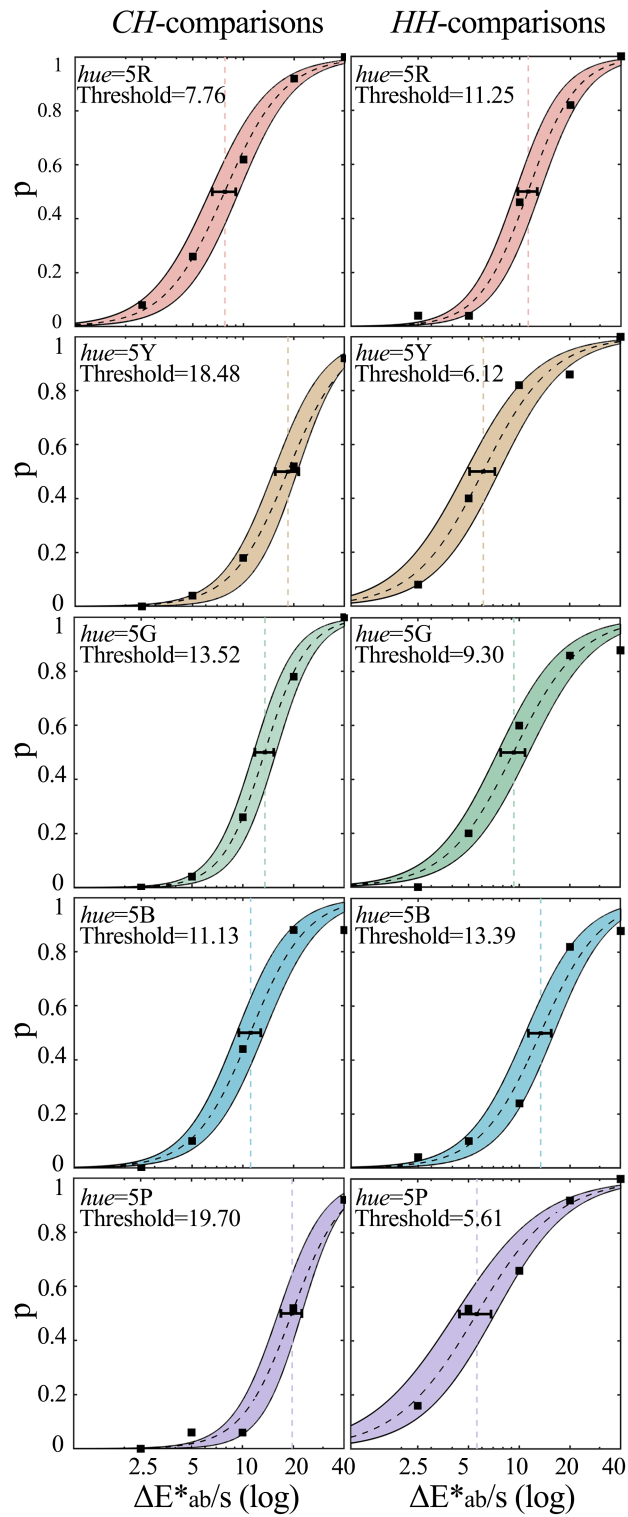


Figure 3.4: Percentage (p) of “the comparison stimulus is faster than the reference stimulus” responses against the transition speed of the comparison stimulus for each of the five hues in the *CH* comparisons (left column) and the *HH* comparisons (right column). The filled squares represent the raw percentages. The colored area represents the 95% psychometric functions, m with the error bars indicating the bootstrap-based 95% confidence intervals.

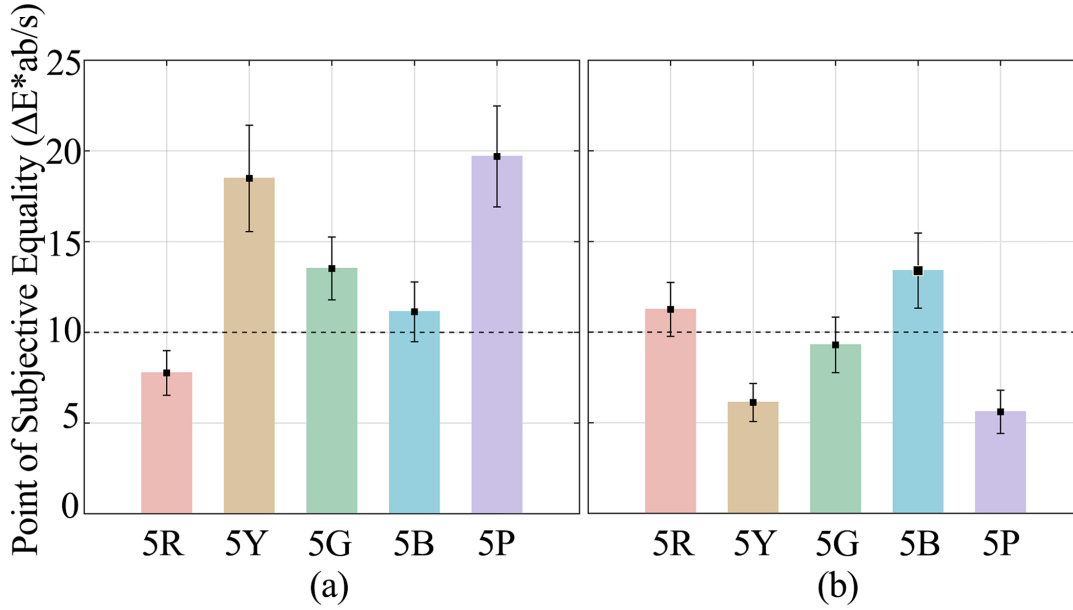


Figure 3.5: Bar chart of the 10 PSEs with their 95% CI for (a) *CH* comparisons and (b) *HH* comparisons.

following notation:

$$PSEs = \begin{vmatrix} 7.76 & 11.25 \\ 18.48 & 6.12 \\ 13.52 & 9.30 \\ 11.13 & 13.39 \\ 19.70 & 5.61 \end{vmatrix} \quad (3.6)$$

where each row represents a base color (in the order of 5R, 5Y, 5G, 5B, and 5P), while the first column represents the *CH* comparisons and the second column the *HH* comparisons. From Figure 3.5, we can conclude that for the *CH* comparisons, the PSE of 5B is not significantly different from the reference speed of $10 \Delta E_{ab}^*/s$, which basically means that the change in chroma and the change in hue around 5B are perceived as equally fast for the same speed in terms of $\Delta E_{ab}^*/s$. At the other base colors, the PSE is significantly different from $10 \Delta E_{ab}^*/s$, which means that the actual speed of the change in chroma needs to be different (in terms of $\Delta E_{ab}^*/s$) from the reference speed of the change in hue in order to let them appear equally fast. Similarly, for the *HH* comparisons, only the PSEs at 5R and 5G are not significantly different from the reference speed, while the others are. For the PSE at 5R, this is not surprising. Actually, the confirmation that a change in hue with a speed of $10 \Delta E_{ab}^*/s$ at 5R is perceived as equally fast as a change in hue with a speed of $10 \Delta E_{ab}^*/s$ at 5R confirms the reliability of our method.

3.3.2. COMPARISON BETWEEN 10 PSEs

When the color transitions would be assessed against the same reference stimulus, their PSE could be directly compared. This was, for example the case for the *HH* comparisons, the PSEs of which could be mutually compared. To relate the PSEs of all color transitions, we expressed them against the same reference stimulus, using the following two assumptions:

1. If speed S_a for transition *A* and speed S_b for transition *B* are perceived as equal, then

transitions with speeds αS_a and αS_b are perceived equal as well (law of proportionality).

2. If the perceived speed of transition A and transition B are equal, and the perceived speed of transition B and transition C are equal, then the perceived speed of A and C are also equal (law of transitivity).

3

Based on these two assumptions, we related the speed of all color transitions to the same reference, being a speed of $10 \Delta E_{ab}^*/s$ in the hue direction around 5R. From the notation in Equation 3.6, we know that a transition speed of $10 \Delta E_{ab}^*/s$ along the hue direction at 5Y was perceived as fast as a transition speed of $(10/6.12) \times 10 = 16.34 \Delta E_{ab}^*/s$ along the hue direction at 5R, and as fast as a transition speed of $18.48 \Delta E_{ab}^*/s$ along the chroma direction at 5Y. Thus, to compare the transition speed along the chroma direction at 5Y with a transition speed along the hue direction at 5R, we need to divide 18.48 by $(10/6.12)$; as a result, a transition speed of $11.31 \Delta E_{ab}^*/s$ along the chroma direction at 5R is assumed to be perceived as fast as a transition speed of $10 \Delta E_{ab}^*/s$ along the hue direction at 5R. Similarly, all the PSE values could be related to the same reference stimulus of a transition speed of $10 \Delta E_{ab}^*/s$ along the hue direction at 5R, which results in

$$PSEs = \begin{vmatrix} 7.76/(10/11.25) & 11.25 \\ 18.48/(10/6.12) & 6.12 \\ 13.52/(10/9.30) & 9.30 \\ 11.13/(10/13.39) & 13.39 \\ 19.70/(10/5.61) & 5.61 \end{vmatrix} = \begin{vmatrix} 8.73 & 11.25 \\ 11.31 & 6.12 \\ 12.57 & 9.30 \\ 14.90 & 13.39 \\ 11.05 & 5.61 \end{vmatrix} \quad (3.7)$$

The PSE values mentioned before can now be directly compared. A value larger than 10 means that the comparison stimulus needs to have a larger speed in order to be perceived equally fast as the reference. Similarly, a value smaller than 10 means that the comparison stimulus needs to have a smaller speed in order to be perceived equally fast as the reference. These relative speeds are depicted in Figure 3.6 for the five base color points and the two directions (in CIELAB) used in the experiment. The length of the arrow in Figure 3.6 represents how the transition speed in $\Delta E_{ab}^*/s$ should be “scaled” in order to make that transition equally fast as the others.

3.3.3. QUESTIONNAIRE ANALYSIS

The answers to the questions we asked at the end of Session 1 and 2 are summarized in Figure 3.7. 65% of the participants indicated that the experiment was difficult (Question 1). They mentioned various reasons, among which the lack of a fixation point to help them focus, and the fact that it was hard to memorize speeds. Only 15% of the participants indicated explicitly that the color change was most visible in their periphery. Some participants (35%) reported that some color transitions were unsmooth at some hue values. All participants except one mentioned that the cap did not influence their judgement; instead, they reported that the cap indeed helped to stop them from looking at the luminaires in the ceiling. Some participants (31%) reported that Session 2 was more difficult because the speed differences were smaller compared to the stimuli in Session 1.

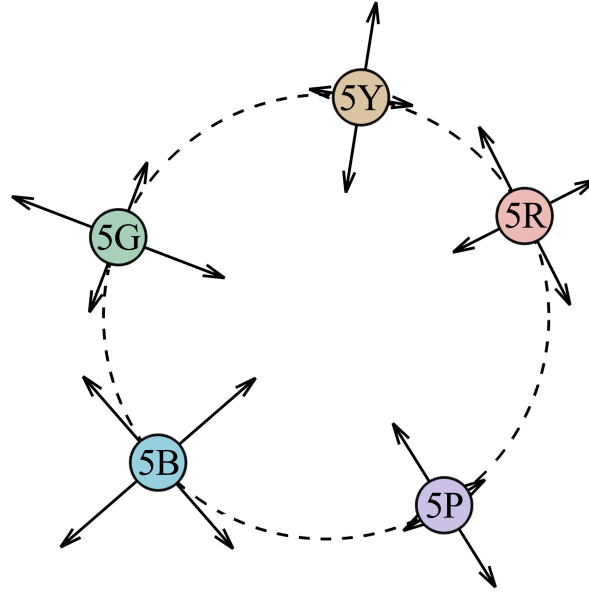


Figure 3.6: Length of the arrows in the figure indicates the PSE of the corresponding color transition to make it equally fast as the reference, being a transition speed of $10 \Delta E_{ab}^*/s$ along the hue direction around the base color 5R (represented by length “1” in this figure).

3.4. COMPARISON WITH AVAILABLE COLOR MODELS

The results presented in Figure 3.5 show that color transitions with the same constant speed in CIELAB are not always perceived as having the same speed. This is no surprise, but it confirms that CIELAB cannot be considered a uniform space to describe the perceived speed of color transitions in the chroma and hue directions. An ideal uniform temporal color model should have equal PSE values independent of the base color and modulation direction of the color transition. A possible way of looking at the uniformity of a color model is to check the standard deviation in PSEs, which is zero in the ideal case.

3.4.1. CALCULATION OF SLOPES IN CIELAB

Despite our assumption that CIELAB would not be temporally uniform, we chose constant $\Delta E_{ab}^*/s$ speeds for this experiment as a starting point. Figure 3.6 already shows how non-uniform CIELAB is for temporal color transitions, and based on these results we made an attempt to improve CIELAB’s uniformity in this respect. This attempt consisted of including a scaling factor in the mutual weighting of $\Delta a^*/s$ and $\Delta b^*/s$ in the calculation of $\Delta E_{ab}^*/s$ as follows:

$$\Delta E_{ab_{imp}}^*/s = \sqrt{(\Delta a^*/s)^2 + (\alpha \Delta b^*/s)^2} \quad (3.8)$$

where $\Delta E_{ab_{imp}}^*/s$ indicates the improved version of $\Delta E_{ab}^*/s$. Due to the triangular shape of the stimuli, a linear fit was performed on the three segments of each temporal transition for the trajectory of a^* and b^* separately. The final $\Delta a^*/s$ $\Delta b^*/s$ and were calculated using the equation

$$Slope = (Slope_1 - Slope_2 + Slope_3)/3 \quad (3.9)$$

where $Slope_1$ covered steps 1–30 of the transition, $Slope_2$ steps 31–90, and $Slope_3$ steps 91–120. Figure 3.8 illustrates how $\Delta a^*/s$ and $\Delta b^*/s$ were obtained.

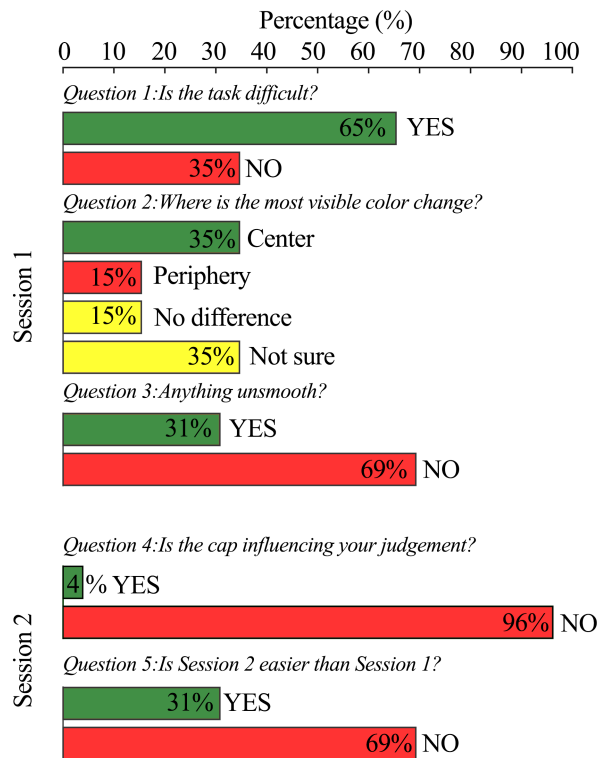


Figure 3.7: Overview of the answers given to the questions at the end of Sessions 1 and 2.

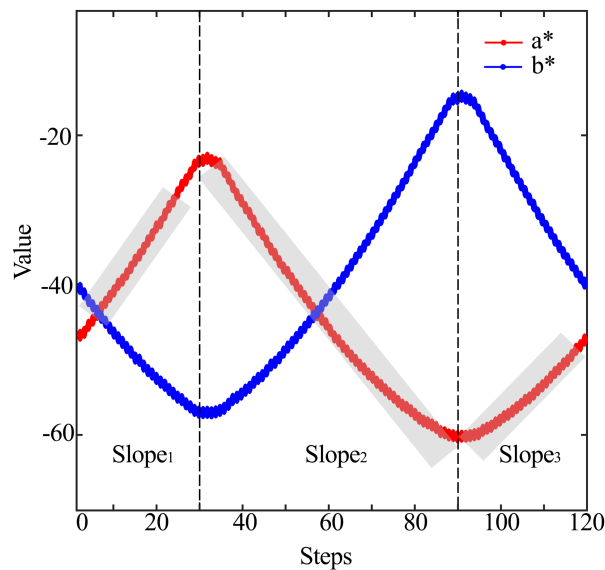


Figure 3.8: Illustration of the segments used for calculating the slope of the a^* and b^* values of the temporal color transition (L^* is constant). $Slope_1$ covers the steps 1–30, while $Slope_2$ and $Slope_3$ cover the steps 31–90 and 91–120, respectively. These segments of steps are shown for the change in a^* values.

Based on these slopes, we computed the standard deviation of the PSEs for different values of the weighting factor α . Figure 3.9 shows that this standard deviation is minimal for a value of $\alpha = 0.404$.

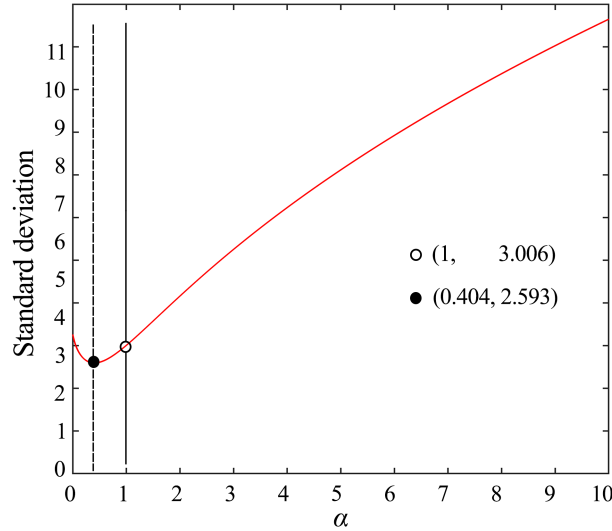


Figure 3.9: Standard deviation of PSEs as a function of α in Equation 3.8. The open circle represents the standard deviation of CIELAB, while the filled circle is the minima found when $\alpha = 0.404$.

3.4.2. CALCULATION OF SLOPES IN LMS CONE FUNDAMENTALS

We designed our dynamic transitions as constant-slope C^* and h changes, which do not correspond to constant-slope L , M , and S changes. To calculate the corresponding L , M , and S cone responses, we used the cone fundamentals of Stockman and Sharpe [133] and the measured spectral data. To minimize the effect of measurement noise, the slopes of $\Delta L/s$, $\Delta M/s$ and $\Delta S/s$ were calculated in a similar way as shown by Equation 3.9. Since neutral white light of 4000 K was regularly shown before, during, and after the stimuli in the experiment, it was chosen as the adaptation condition; in reality, we expect that observers' state of adaptation was influenced by the dynamic stimuli, but we don't have a way to predict this variation. Thus, the L , M , and S values at each step were normalized by dividing them by the L , M , and S values of 4000 K, which is equal to multiplying by 3.28, 4.11, and 11.22, respectively. However, this normalization step does not influence the calculation of slopes.

Using the same concept of minimizing the standard deviation of PSEs, a scaled LMS space with parameters α and β could be constructed as follows:

$$\Delta LMS/s = \sqrt{(\Delta L/s)^2 + \alpha(\Delta M/s)^2 + \beta(\Delta S/s)^2} \quad (3.10)$$

The minimum in standard deviation was found when α and β were both zero, which means that the M and S responses were eliminated. Eliminating these responses may then give the best fit, but it is intuitively unsatisfying. Thus, based on the physiologically based DKL space [12], which has proven its value in the literature, the following equation was constructed to minimize the standard deviation of the PSEs:

$$\Delta DKL = \sqrt{[\Delta(L - M)/s]^2 + \alpha \{ \Delta[S - (L + M)]/s \}^2} \quad (3.11)$$

Since the $L + M$ channel usually represents luminance, which in our case remains constant, the slope of the $L + M$ channel is ideally zero. Using the same technique as described previously, we now found a minimum for the standard deviation in PSE when $\alpha = 0.02$ (as shown in Figure 3.10).

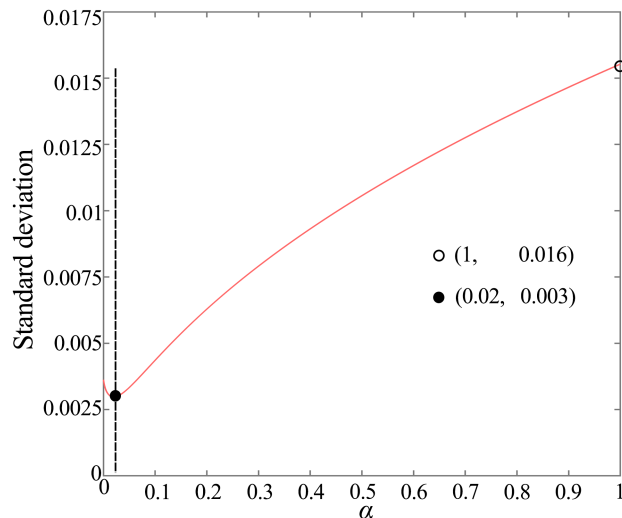


Figure 3.10: Standard deviation of PSEs as a function of α in Equation 3.11. The open circle represents the standard deviation of DKL, while the filled circle is the minima found when $\alpha = 0.02$.

3.5. DISCUSSION

3.5.1. OPTIMIZATION OF COLOR SPACES

Our results demonstrate that more uniformity can be achieved by improving available color spaces. Specifically, the CIELAB-based improvement suggests that $\Delta b^*/s$ contributes less to speed perception than $\Delta a^*/s$, but the resulting reduction in the standard deviation of the PSEs is very small. The improved DKL-based space suggests that $\Delta[S - (L + M)]/S$ hardly contributes to speed perception compared to $\Delta(L - M)/s$. In this way, the standard deviation in PSEs is reduced by almost a factor of 6. It is important to notice that the small weight ($\alpha = 0.02$) for the $S - (L + M)$ channel may suggest that the $L - M$ channel dominates the speed perception of chromatic change. Although there is no literature for the direct comparison with our results, several papers indicated the “weak” contribution of the $S - (L + M)$ channel in speed perception. For example, Dougherty, Press, and Wandell [17] proposed an equation that combines the luminance, red–green, and blue–yellow opponent mechanisms to describe the mechanism of moving color responses. They assigned the largest weight to the luminance channel ($L + M$), followed by the weight in the red–green $L - M$ channel and then the blue–yellow [$S - (L + M)$] channel. Nguyen-Tri and Faubert [101] performed a speed-matching experiment using drifting chromatic gratings. They found that the $S - (L + M)$ post-receptoral mechanism does not appear to contribute significantly to determining the perceived speed of chromatic motion. McKeefry and Burton [90], however, did not agree with the opinion that S-cone signals are not as effective in signaling stimulus motion. Their results suggest that the perception of speed is based on $L - M$ and $S - (L + M)$ cone opponent processing.

Figure 3.11 shows the bar charts of the 10 PSEs in the four spaces: CIELAB and improved CIELAB [Figure 3.11(a)], DKL and improved DKL [Figure 3.11(b)]. It is not intuitively meaningful to compare these two metrics, as they are of different magnitude. The fact that the optimized DKL space has reduced the standard deviation of PSEs to a larger extent than optimized CIELAB might indicate that DKL is more promising to work with in future. However, that this space implies that the $S - (L + M)$ responses can be largely neglected for

speed perception of color transitions remains worrisome. It might as well be too simplistic to just target the minimum of the standard deviation of PSEs to improve the uniformity of a color space for temporal transitions. Hence, to gain more insight into how to build a color space with global uniformity for speed perception of temporal color transitions, more data collection and more extensive modeling are needed.

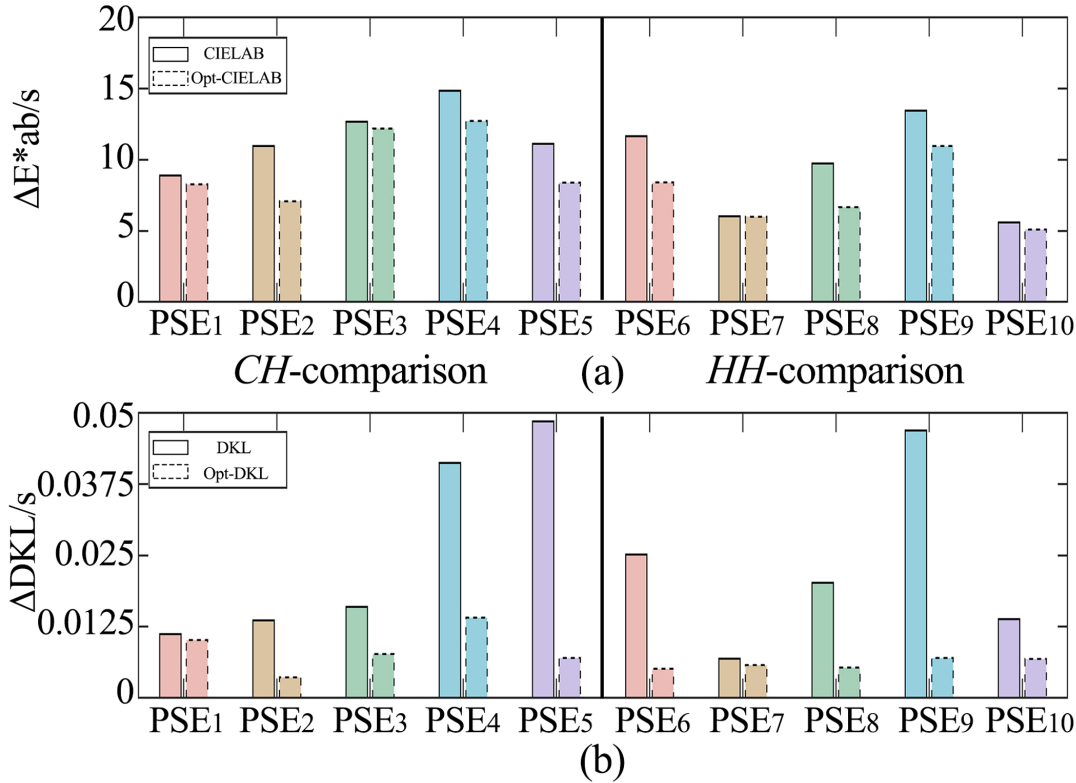


Figure 3.11: Bar charts of the 10 PSEs expressed in different color spaces: (a) CIELAB and improved CIELAB; (b) DKL and improved DKL.

3.5.2. LIMITATIONS AND FUTURE WORK

Our stimuli temporally varied in C^* and h , keeping the luminance constant. For the design of these stimuli, we kept the luminance constant for an averaged observer. It is, however, generally known that luminance perception varies over individuals, and consequently, our stimuli might not have been fully isoluminant for all our participants. Nonetheless, we estimate the effect of individually tuned isoluminance to be small for this particular experiment. First of all, various researchers have found that adults are more sensitive to chromatic than luminance contrast for temporal frequencies below 4–5 Hz [15]. In this experiment specifically, the frequency of the temporal transition was about 0.33 Hz, and at this frequency the data of [15] indicate that adults are at least four times more sensitive to chromatic change than to luminance change. Second, if we simulate possible variations in the $(L + M)$ channel of our stimuli using the CIE 2006 Physiological Observer (CIEPO06) [69], we found that these variations have a slope close to zero, and at least five times smaller than the slope in any of the chromatic channels [i.e., either $L - M$ or $S - (L + M)$]. Still, in our next experiments to extend our data, we will evaluate the effect of including individual

isoluminance in the design of our stimuli.

In the current experimental design, lightness changes were not included. Since our long-term goal is to develop a temporal uniform color space, comparing the perceived speed of lightness change with chromaticity change is important. In addition, as mentioned before we omitted the *CC*-comparisons in this study, but they clearly are relevant to extend the set of stimuli for the modeling or to validate the improved color space for temporal color transitions. In the future, lightness changes and more chroma levels will also be included in the experimental design.

Since the task of comparing multiple pairs of sequential color transitions is quite time-consuming, the number of repetitions per participants had to be limited and we were not able to fit a psychometric function per participant. In contrast, the responses were averaged over participants to fit one psychometric function for each hue in the *CH* comparisons and each hue in the *HH* comparisons. We already know that fitting a psychophysical curve across observers may introduce between-observer noise. Thus, if time would allow, an ideal case would be fitting a psychometric function per observer. This would require a more efficient psychophysical method or a reduction in the number of experimental conditions, where the latter is undesirable as well.

We intentionally chose for full room setting instead of the often-used 2° or 10° stimuli because temporal dynamics in artificial lighting are more realistically represented with a room-size field of view. At the same time, this choice introduces uncertainty about the direction of the observer's gaze and the effects of peripheral vision. For the observers sitting at the predefined position, our stimuli resulted in a 101° horizontal and roughly 77° vertical visual angle. Research has shown that both spatial vision [39] and temporal vision [105] vary across this eccentricity, but our data does not allow a comparison of temporal color perception in central or peripheral vision. However, in lighting applications, central and peripheral vision will always mix, so a generally suitable temporal color space is needed. Creating models that take eccentricity into account would require new data in the future.

We chose to use a wall illumination of 4000 K as an intermediate light in an attempt to continuously reset the participants' state of adaptation. Still, participants constantly adapted during the experiment to the varying light environment they were immersed in. This may have affected the experimental results, though most probably in a systematic way. This surely impacts the modeling, since color spaces assume adaption to a fixed illuminant. Due to the lack of a good model for temporal chromatic adaptation (which would change *L*, *M*, and *S* reference over time), the calculation of the cone responses to model speed perception can be further improved.

Finally, it should be noted that we might have accumulated error in our calculations to compare the perceived speed of various color transitions with one fixed reference. The accumulated error can be best estimated from the first row of the result summary in Equation 3.7. In our experiment, we directly measured that the transition speed along the chroma direction at 5R needed to be $7.76 \Delta E_{ab}^*/s$ in order to be perceived as fast as a transition speed of $10 \Delta E_{ab}^*/s$ along the hue direction at 5R. Using the calculation towards the same reference stimulus, we estimate a transition speed of $8.73 \Delta E_{ab}^*/s$ along the chroma direction at 5R to be perceived as fast as a transition speed of $10 \Delta E_{ab}^*/s$ along the hue direction at 5R. The deviation results from the fact that the *HH* comparison with both the comparison and reference stimulus at 5R yielded a PSE of 11.25 instead of $10 \Delta E_{ab}^*/s$. It is worth pointing out that $\Delta a^*/s$ and $\Delta b^*/s$ are linear for *C** transitions, but sinusoidal for *h* transitions, which

also has a drawback in the calculation of the slopes according to Figure 3.8.

Based on the just-mentioned limitations, we see multiple options for further research. Perception of dynamic color has many dimensions, such as three dimensions to define a base point in a color space, three dimensions of the modulation direction, and the related periodicity and waveform. In the future, more base colors (hues) and modulation directions (not only cardinal directions) will be needed to more accurately model a temporally uniform color space. Instead of defining the related stimuli in CIELAB, we see added value in defining them directly in cone-contrast space. In addition, there are more ways of generating the shape of the stimuli. Rather than giving all stimuli a fixed time period and changing the speed by changing the amplitude of the chroma or hue, one could keep the chroma or hue amplitude constant and change the speed by changing the time interval. The disadvantage of this choice, though, may be that observers use the additional time cue to judge the perceived speed of the color transition. Finally, it could be advantageous to use different JNDs instead of PSEs to distinguish perceived speeds, so they can be further used to build the temporal sensitivity function.

3.6. CONCLUSIONS

In this study, the perceived speed of periodic color transitions in chroma and hue was measured, starting from the five principal Munsell hues. The results indicate that available color spaces are not able to predict our speed perception. It is not so surprising that CIELAB is not perceptually uniform for temporal color transitions, as it was designed for describing spatial color differences. Color spaces built on cone responses were also insufficiently accurate to predict uniform speed perception of color transitions. Simple modifications of these spaces improve the uniformity in speed perception, but we foresee more extensive experimentation to build a more accurate model.

4

Assessing the Temporal Uniformity of CIELAB Hue Angle

In **Chapter 3**, we found that $\Delta E_{ab}^*/s$ in CIELAB is not suitable for describing the perceived speed of temporal color changes in full-room illumination. Two hue transitions with the same physical speed of change, in terms of $\Delta E_{ab}^*/s$, were not perceived to change at the same speed. This is not really surprising, since CIELAB was not designed to characterize the perception of temporal color transitions in illumination. In this study, we further investigate the temporal uniformity of CIELAB. The stimuli were presented in a square of 4.3° visual angle surrounded by a 4000 K adapting field, similar to the viewing condition for which CIELAB was designed (i.e., where color stimuli are presented on-axis surrounded by a static adaptation field). The human observers viewed pairs of temporal color transitions which were presented sequentially, and were asked to select the one that appeared to change faster. The results confirmed that under these conditions CIELAB was also not temporally uniform. We present preliminary attempts to improve the temporal uniformity for both CIELAB and cone-excitation spaces (i.e., LMS and DKL (Derrington–Krauskopf–Lennie [12])).

This chapter is copied with (slight) adaptations from Journal of the Optical Society of America A, Volume 37, Number 4, 521-528 (2020) as: "Assessing the Temporal Uniformity of CIELAB Hue Angle", Xiangzhen Kong, Minchen Wei, Michael J. Murdoch, Ingrid Vogels, and Ingrid Heynderickx [59].

4.1. INTRODUCTION

Human perception of a color stimulus depends not only on the spectral composition of the stimulus, but also on its surrounding [2, 9]. In order to characterize perceived color differences, various color spaces have been developed based on the results of psychophysical experiments. In these experiments, human observers performed a color matching task or scaled color differences between pairs of static stimuli which were always presented simultaneously. Ideally, a useful color space should be perceptually uniform in order to describe the perception of color differences. However, MacAdam [84] already demonstrated the spatial non-uniformity of the CIE 1931 chromaticity diagram by deriving ellipses showing the ranges of equivalent color matches (known as the MacAdam ellipses).

4

With the development of LED technology, the color and intensity of lighting systems can be easily adjusted to create appealing light atmospheres [145, 40, 97, 79]. We can adjust the illumination in the entire space or only change the illumination in a small area, with the former changing the stimulus for a large field of view (FOV) and the latter changing the stimulus for a small FOV (e.g., a small stage area illuminated with a projection light). Moreover, the colors and intensities of displays or media walls can also be dynamically changed to enhance the viewing experience of observers [115]. Both the ambient illumination and characteristics of a stimulus can change dynamically. Literature has shown that the rate of change in illumination influences whether we perceive the (static) material world as changing or stable [10, 1]. Though interesting and important itself, for the purpose of our study—to investigate the perceived speed of a light transition for a homogeneous stimulus—we focus on the speed perception of a dynamic stimulus under a constant adapting condition.

To the best of our knowledge, very few studies have investigated whether color spaces that are known to be uniform for spatial color differences can accurately characterize perceived differences of temporal changes in color (i.e., temporal uniformity). In [121], the observers were asked to adjust the step size along the three dimensions in CIELAB LCh color space, so that the temporal transitions appeared smooth. A three-primary (i.e., red, green, and blue) LED light source was used to produce the stimulus, which was projected on a wall with a dark surround, and the observers were seated three meters away from the wall. It was found that the threshold to produce a smooth transition (i.e., the maximum color difference between two successive colors for producing a smooth transition) along the chroma and hue directions was around 10 times the threshold along the lightness direction. In **Chapter 3**, the observers compared the perceived speed of temporal changes along chroma and hue directions. The experiment was designed to find the point of subjective equality (PSE) of speed between different color transitions. The full-room sized stimuli were sequentially presented with a static 4000 K illumination in between. It was found that $\Delta E_{ab}^*/s$ failed to characterize the perceived speed of change. These two studies employed similar viewing conditions; they used large field-of-view illumination, so that the observers used both central and peripheral vision to evaluate the color perception of the stimuli [39, 135]. Such setups, however, are different from the typical viewing conditions for which CIELAB was designed, i.e., where color stimuli are presented on-axis surrounded by a static adaptation field.

Therefore, this study was specifically designed to evaluate the temporal uniformity of CIELAB for stimuli surrounded by a static adaptation field. Apart from that, this study is similar to the one in **Chapter 3** in order to allow comparison of the results. Pairs of stimuli

(FOV \approx 4.3°) were temporally modulated in hue angle around different base colors and were sequentially presented in a static adaptation field. The observers compared the perceived speed of the temporal change between the two stimuli in each pair and selected the one that appeared to change faster. To allow a direct comparison, the luminance, hue angles, and chroma levels of the stimuli, and the temporal modulations (i.e., resolution and waveform) were the same as those in **Chapter 3**. The luminance and correlated color temperature (CCT) of the adaptation field was equal to the 4000 K stimulus used between the alternating stimuli in our previous study.

4.2. METHODS

The experiment was carried out in the Color and Illumination Laboratory at The Hong Kong Polytechnic University. The procedure and protocol of the experiment were approved by the Institutional Review Board at The Hong Kong Polytechnic University.

4.2.1. APPARATUS

A viewing booth, with dimensions of 60 cm (height) \times 60 cm (width) \times 60 cm (depth), was built for this experiment. The interior of the booth was painted using Munsell N7 spectrally neutral paint. A 4.5 cm \times 4.5 cm square opening was cut out at the center of the back panel, as shown in Figure 4.1, and was covered with a diffuser. A THOUSLITE LEDCube consisting of 11 channels with the peak wavelength between 400 and 700 nm was placed on top of the viewing booth to provide a uniform illumination. An ARRI L5-C LED fixture was installed on a tripod behind the viewing booth and it was carefully adjusted to uniformly illuminate the diffuser in the back panel.

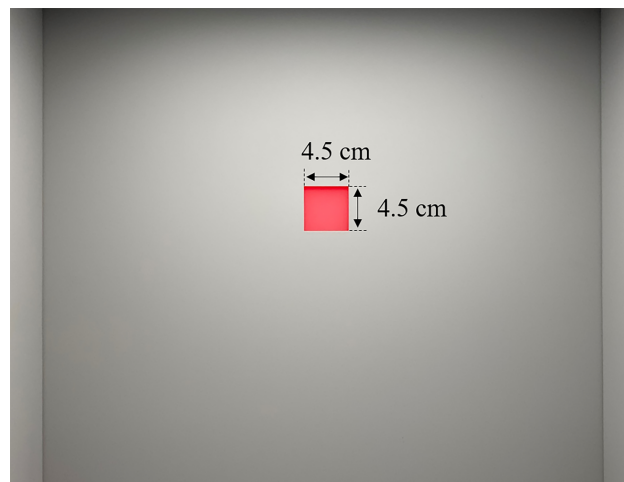


Figure 4.1: Photograph of the experimental setup captured at the observer's eye position. An LEDCube was placed above the booth to produce surrounding illumination in the booth as the adaptation field. An ARRI L5-C LED fixture was fixed on a tripod and placed behind the viewing booth. It produced illumination to the opening (4.5 cm \times 4.5 cm) as the stimulus.

A chin rest was mounted in front of the viewing booth, centered such that the opening on the back panel was at the observer's eye height, subtending a 4.3° \times 4.3° FOV. The top part of the front opening was partially covered using black felt to prevent the observer from seeing the LEDCube directly.

4.2.2. ADAPTATION CONDITION AND STIMULI

The intensities of the 11 channels of the LEDCube were carefully adjusted, so that the luminance and CCT of the adaptation field (as measured at the opening in the back panel using a calibrated Labsphere reflectance standard and a JETI Specbos 1211UV spectroradiometer) were similar to the 4000 K stimulus in **Chapter 3**. In the new setup we measured a CCT of 4080 K with a D_{uv} of +0.003 and at a luminance $L_{w,10}$ (i.e., the luminance of a perfect reflectance standard) of 210 cd/m².

The intensities of the four channels (i.e., red, green, blue, and white) in the ARRI L5-C fixture were controlled, with a resolution of 16 bit, via Digital Multiplex (DMX). A USB-to-DMX interface (i.e., the DMXKing ultraDMX Micro) was used to control the fixture through a desktop. The experiment only used three channels (i.e., red, green, and blue), which were calibrated using the JETI spectroradiometer under the adaptation condition. The calibration was performed by setting a series of input signals to the three channels and measuring the tristimulus values of the produced stimuli. A good linearity was found between the input signals and the tristimulus values of the produced stimuli, so that the relationship between them could be characterized using Equation 4.1 with its nine coefficients being derived from the measurements. Subsequently, Equation 4.1 was used to calculate the corresponding input signals to the three channels for producing the desired stimuli. Specifically, for our experimental setup, Equation 4.1 becomes

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 410.693 & 136.748 & 255.655 \\ 185.537 & 518.630 & 141.704 \\ 1.325 & 77.412 & 1445.863 \end{bmatrix} \begin{bmatrix} R/R_{max} \\ G/G_{max} \\ B/B_{max} \end{bmatrix} + \begin{bmatrix} 27.968 \\ 28.929 \\ 18.790 \end{bmatrix} \quad (4.1)$$

where XYZ are the tristimulus values of a stimulus being calculated using the CIE 1976 10° color matching functions (note: this included both the stimulus produced by the ARRI fixture and the light produced by the adapting condition that was reflected by the diffuser at the opening); R , G , and B are the input signals to the three channels; R_{max} , G_{max} , and B_{max} are the maximum input signals (i.e., $2^{16}-1 = 65535$); 27.968, 28.929, and 18.790 represent tristimulus values of the reflected light from the adapting condition when the three input signals were set to zeros.

To make the results comparable with **Chapter 3**, the same stimuli were adopted. The stimuli were periodic temporal color transitions around a base color point in the hue direction of CIELAB LCh color space, with the measured adaptation field used as the white point. The five base color points were set at hue angles of 27°, 81°, 159°, 221°, and 302°, with a chroma C^* of 60 and a lightness L^* of 90. Four color transition speeds (CTSs), i.e., 5, 10, 20, and 40 $\Delta E_{ab}^*/s$, were used, and the corresponding CTSs in hue were calculated using the following equation:

$$\Delta H^* = \arccos \left(1 - \frac{\Delta E_{ab}^{*2}}{2 \times C^{*2}} \right) \quad (4.2)$$

Each color transition consisted of 120 discrete steps, and was presented as a triangular-shaped waveform to allow an easier judgement of the speed of the transition. The color changed every 0.025(= Δt)s over a full period of 3(= 120×0.025) s. To avoid an abrupt change that might influence the perception of speed, as suggested in [121], the sharp edges of the triangular wave were smoothed (see the regions shaded in gray in Figure 4.2). Specifically, Steps 1 to 27, Steps 34 to 87, and Steps 94 to 120 had a fixed step size S , while the smoothed region had half of the step size as its previous adjacent step, resulting in a step size of

Table 4.1: Hue angle ranges for the transitions around the different base colors with different speeds

Speed $\Delta E_{ab}^*/s$	Base Color (Hue Angle) and Its Corresponding Amplitude				
	Red (27°)	Yellow (81°)	Green (159°)	Blue (221°)	Purple (302°)
5	$27^\circ \pm 3.33^\circ$	$81^\circ \pm 3.33^\circ$	$159^\circ \pm 3.33^\circ$	$221^\circ \pm 3.33^\circ$	$302^\circ \pm 3.33^\circ$
10	$27^\circ \pm 6.66^\circ$	$81^\circ \pm 6.66^\circ$	$159^\circ \pm 6.66^\circ$	$221^\circ \pm 6.66^\circ$	$302^\circ \pm 6.66^\circ$
20	$27^\circ \pm 13.37^\circ$	$81^\circ \pm 13.37^\circ$	$159^\circ \pm 13.37^\circ$	$221^\circ \pm 13.37^\circ$	$302^\circ \pm 13.37^\circ$
40	$27^\circ \pm 27.14^\circ$	$81^\circ \pm 27.14^\circ$	$159^\circ \pm 27.14^\circ$	$221^\circ \pm 27.14^\circ$	$302^\circ \pm 27.14^\circ$

$\frac{1}{2}S$, $\frac{1}{4}S$, and $\frac{1}{8}S$, respectively. A more detailed description of the stimulus modulation was presented in **Chapter 3**. The CTS was defined as the speed of the unsmoothed region of the transition (i.e., $S/\Delta t$), and the maximum hue angle A that deviated from the base color hue was calculated using

$$A = (27 \times S) + \left(\frac{1}{2} \times S + \frac{1}{4} \times S + \frac{1}{8} \times S \right) \quad (4.3)$$

where $S = \Delta H^* \times \Delta t$.

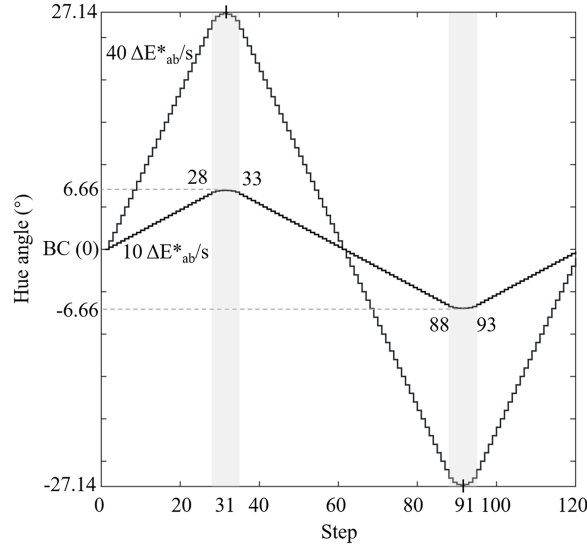


Figure 4.2: Illustration of the stimulus modulation in the hue direction of CIELAB LCh color space, with a speed of $10 \Delta E_{ab}^*/s$ and $40 \Delta E_{ab}^*/s$. The abscissa represents the step number, while the ordinate represents the hue angle deviation from the base color. Steps 1 to 27, Steps 34 to 87, and Steps 94 to 120 had a fixed step size S , while the smoothed area (i.e., Steps 28 to 33 and Steps 88 to 93, the regions shaded in gray) had half of the step size as its previous adjacent step. For a speed of $10 \Delta E_{ab}^*/s$ and $40 \Delta E_{ab}^*/s$, the maximum hue angle deviation from the base color is 6.66° and 27.14° , respectively.

Given the above, the ranges of the hue angles can be calculated based on the speed $\Delta E_{ab}^*/s$, as shown in Table 4.1. Subsequently, the intermediate hue angles of all stimuli can be calculated, and transformed to the input signals (i.e., R , G , and B) of the fixture, using Equation 4.1.

The JETI spectroradiometer was used to measure the spectral power distributions of all the produced stimuli (i.e., Steps 31 to 91 of the triangular waveform in Figure 4.2). Figure 4.3 shows the comparison of the hue angles between the produced stimuli and the designed stimuli. It can be observed that the stimuli in the red region, on average, have the largest

hue angle deviation (about 2.5°), which was likely due to the instability of the red LEDs. The deviations in the other regions were smaller with an average hue angle difference about 1° , except the deviation of the transition around 81° with a speed of $40 \Delta E_{ab}^*/s$, which was so large that it overlapped with the transition around red.

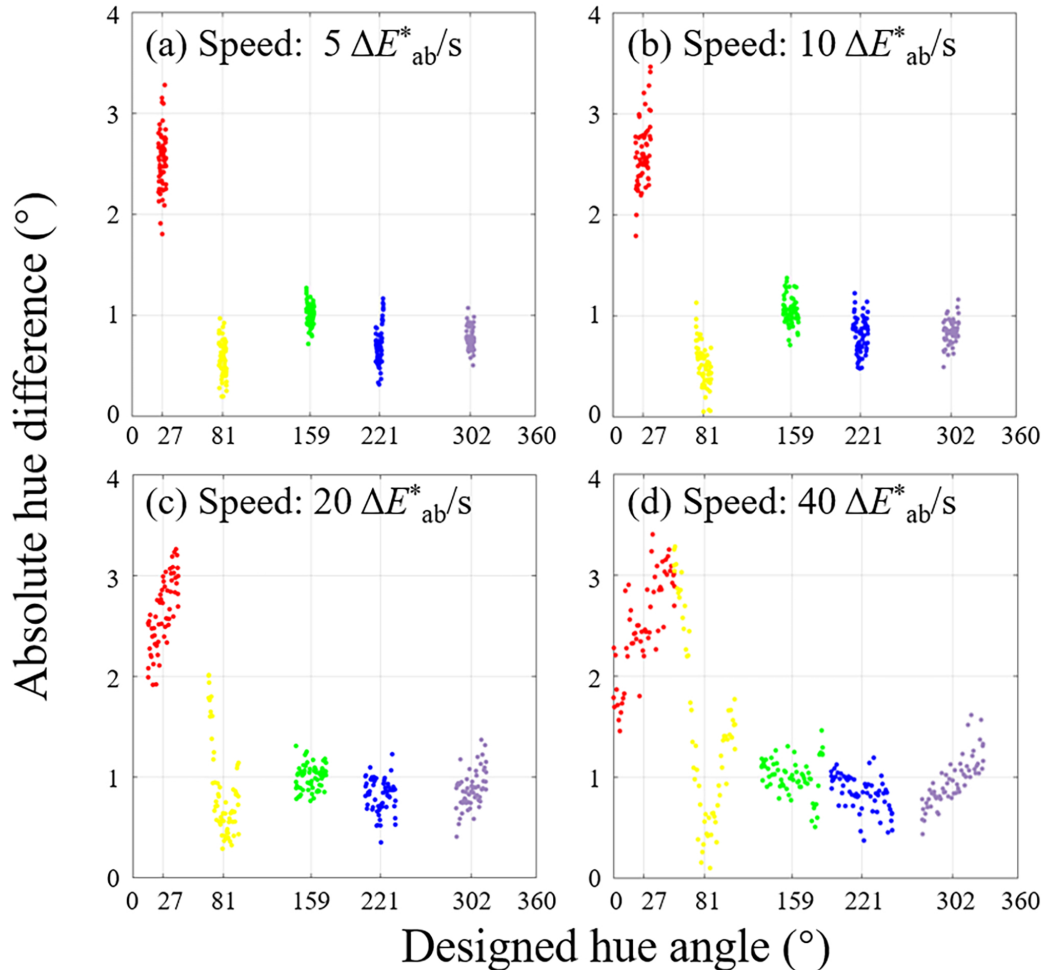


Figure 4.3: Comparison between the designed and measured hue angles of the stimuli included in the color transitions with different speeds. (a) $5\Delta E_{ab}^*/s$, (b) $10 \Delta E_{ab}^*/s$, (c) $20 \Delta E_{ab}^*/s$, (d) $40 \Delta E_{ab}^*/s$. The y axis is the absolute difference between the measured and the designed hue angles while the x axis is the designed hue angle.

A final note of attention is that a DMX needs 22 ms (i.e., 44 Hz) to transfer a message of 512 bytes [6]. Given the smaller number of bytes and the hardware used in the experiment, efforts were made to precisely hold the refresh rate at 40 Hz by including delays through a computer program. Such implementations were used in **Chapter 3** and found reliable [61, 96].

4.2.3. EXPERIMENTAL CONDITIONS

As mentioned above, to allow a direct comparison with our work in **Chapter 3**, we used the same reference stimulus to which the other color transitions had to be assessed, namely a color transition around the red base color at a speed of $10 \Delta E_{ab}^*/s$. In the pilot experiments,

we found that the color transitions around the green base color yielded results very different from the other stimuli. We speculated that such a difference was due to the fact that the greenish stimuli varied around an opponent color from the (red) reference transition. Since these two opponent transitions were successively presented without an intermediate adaptation stimulus (as used in **Chapter 3**), we anticipated biased PSE results. Therefore, we employed an additional reference stimulus, i.e., the transition around the base color of green at a speed of $10 \Delta E_{ab}^*/s$, to which transitions around the base color of purple were assessed. This pair of colors was carefully selected, since the transitions around the base color of purple had the most different PSE from green (i.e., hue angle of 159°) in **Chapter 3**.

When using the method of constant stimuli, the stimulus levels, in our case speeds, should be carefully selected. Both pilot experiments and our work in **Chapter 3** suggested that the lowest speed ($2.5 \Delta E_{ab}^*/s$) and highest speed ($40 \Delta E_{ab}^*/s$) were perceived obviously different from the reference speed of $10 \Delta E_{ab}^*/s$. As found in **Chapter 3**, only nine of the 130 comparisons at $2.5 \Delta E_{ab}^*/s$ were judged faster than the reference speed of $10 \Delta E_{ab}^*/s$, and only eight of the 130 comparisons at $40 \Delta E_{ab}^*/s$ were judged slower than the reference speed. Therefore, these two speeds were not expected to affect the calculation of PSE values. As a result, in the current experiment, we discarded the stimuli with a speed of $2.5 \Delta E_{ab}^*/s$, and only included the stimuli with a speed of $40 \Delta E_{ab}^*/s$.

In summary, each condition included two color transitions—a reference stimulus and a comparison stimulus—which were sequentially presented. The reference stimulus was always the color transition around the base color of red at a speed of $10 \Delta E_{ab}^*/s$; the comparison stimuli were designed to include transitions around five base colors (i.e., red, yellow, green, blue, and purple) and four transition speeds (i.e., 5, 10, 20, and $40 \Delta E_{ab}^*/s$). An additional reference stimulus with the base color of green at a speed of $10 \Delta E_{ab}^*/s$ was used for the comparison stimuli with the base color of purple at the four transition speeds. In total, 24 conditions (i.e., 5 hues \times 4 speeds + 1 hue \times 4 speeds) were included in the experiment.

4.2.4. OBSERVERS

Twenty-one observers (14 males and 7 females) between 20 and 30 years of age (mean = 22.8, standard deviation = 2.6) participated in the experiment. All of them had a normal color vision, as tested using the Ishihara Color Vision Test.

4.2.5. PROCEDURE

When the observer arrived in the lab, the Ishihara Color Vision Test was carried out, followed by the collection of demographic information such as age, gender, and whether the observer had any self-reported vision problems. He/she was then seated in front the viewing booth and was asked to adjust the height of the chair, so that he/she could comfortably fix his/her chin on the rest. The adaptation field was presented in the booth, while the general illumination in the rest of the experimental room was switched off. The experimenter described the experimental procedure to the observer. Then, the observer evaluated four practice trials to get familiar with the procedure and task. These four practice conditions were identical to all observers and were carefully selected, so that they were representative for the comparisons of colors and speeds occurring during the rest of the experiment (i.e., red at $40 \Delta E_{ab}^*/s$ versus red at $5 \Delta E_{ab}^*/s$; green at $10 \Delta E_{ab}^*/s$ versus red at $40 \Delta E_{ab}^*/s$; red at $10 \Delta E_{ab}^*/s$ versus red at $5 \Delta E_{ab}^*/s$; red at $10 \Delta E_{ab}^*/s$ versus purple at $5 \Delta E_{ab}^*/s$).

During the experiment, computer-generated verbal instructions were used, similar to

Chapter 3. These instructions were: **V1** “First stimulus”; **V2** “Second stimulus”; **V3** “Which stimulus appears faster?”; **V4** “Invalid input, please input again”; **V5** “Input recorded”; and **V6** “Congratulations, you finished the experiment. Thank you for your participation.” When the reference and comparison stimulus were alternating, the instructions **V1** and **V2** were played. After two alternations, the instruction **V3** prompted the observer to make a forced choice about which stimulus appeared faster and to input the judgment using a keypad.

One main difference in the procedure compared to the one used in **Chapter 3** concerned the adaptation stimulus played in between the reference and comparison stimulus. It was reported in **Chapter 3** that 65% of the observers found the experimental task difficult. One common reason was that it was hard to memorize the perceived speed. Because the intermediate adaptation field was presented for 10 s, there was a delay of 25 s between the end of the entire presentation and the end of the first stimulus. To allow an easier comparison between the two stimuli, the presentation of the stimuli was changed, as shown in Figure 4.4. Now, the two stimuli were alternated back and forth twice, with a dark period (i.e., ARRI L5-C was turned off) of 0.5 s between each alternation.

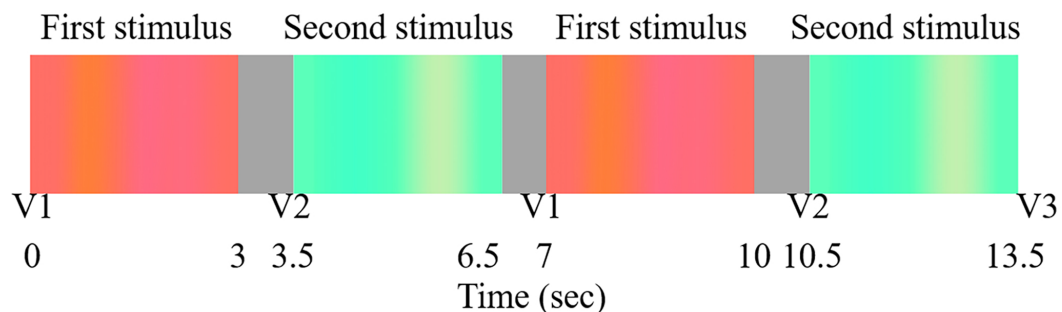


Figure 4.4: Illustration of the procedure for each comparison. Each stimulus in a pair was presented twice. V1, V2, and V3 are the voice instructions: “First stimulus,” “Second stimulus,” “Which stimulus appears faster?”

Each observer evaluated 24 conditions twice, with the order of the two stimuli in each condition being counterbalanced. The 48 comparisons were presented in a random order. The experiment took around 20 min for each observer.

4.3. RESULTS

4.3.1. POSSIBLE BIAS

As for all stimuli the order was counterbalanced in the repetition, the first stimulus should be selected as often as the second stimulus when there is no response bias. Of the 1008 judgments in total (i.e., 4 speeds \times 6 conditions \times 2 repetitions \times 21 observers), the first stimulus was selected 543 times (53.9%) and the second stimulus 465 times (46.1%), which was significantly different from 50% as tested using a Chi-square goodness-of-fit test ($\chi^2(1, N = 1008) = 6.04, p\text{-value} = 0.014$). Additional tests were performed on each condition to investigate whether the presentation order and the selection of the stimulus were independent to each other using a Chi-square test of independence. Only the condition with the comparison stimulus around the base color of purple at a speed of $5 \Delta E_{ab}^*/s$ was found to have an association between the presentation order and the selection of the stimulus ($\chi^2(1, N = 42) = 9.55, p\text{-value} = 0.002$). For the “null condition,” in which the comparison stimulus was identical to the reference stimulus having the base color of red at a speed of $10 \Delta E_{ab}^*/s$,

the first stimulus was selected 23 times (54.8%) and the second stimulus 19 times (45.2%), which was not significantly different from 50% as tested using a Chi-square goodness-of-fit test ($\chi^2(1, N = 42) = 0.38, p\text{-value} = 0.54$). In addition, among the 42 judgments on the conditions, in which the comparison stimulus had a speed of $40 \Delta E_{ab}^*/s$, the reference stimulus was only selected six times. These analyses demonstrated that the experimental setup and procedure can be considered reliable, and that possible effects of bias compensations were considered minor.

4.3.2. CALCULATION OF PSEs AND CONFIDENCE INTERVALS

For each base color, the percentage of judgements “the comparison stimulus appeared faster than the reference stimulus” is plotted as a function of the speed of the comparison stimulus in Figure 4.5. For the 24 conditions that the comparison stimulus was not identical to the reference stimulus, a Chi-square goodness-of-fit test was employed to test whether the judgment was significantly different from random chance (i.e., 50%). Five conditions, labeled as filled circles in Figure 4.5, were found not significantly different from a random chance.

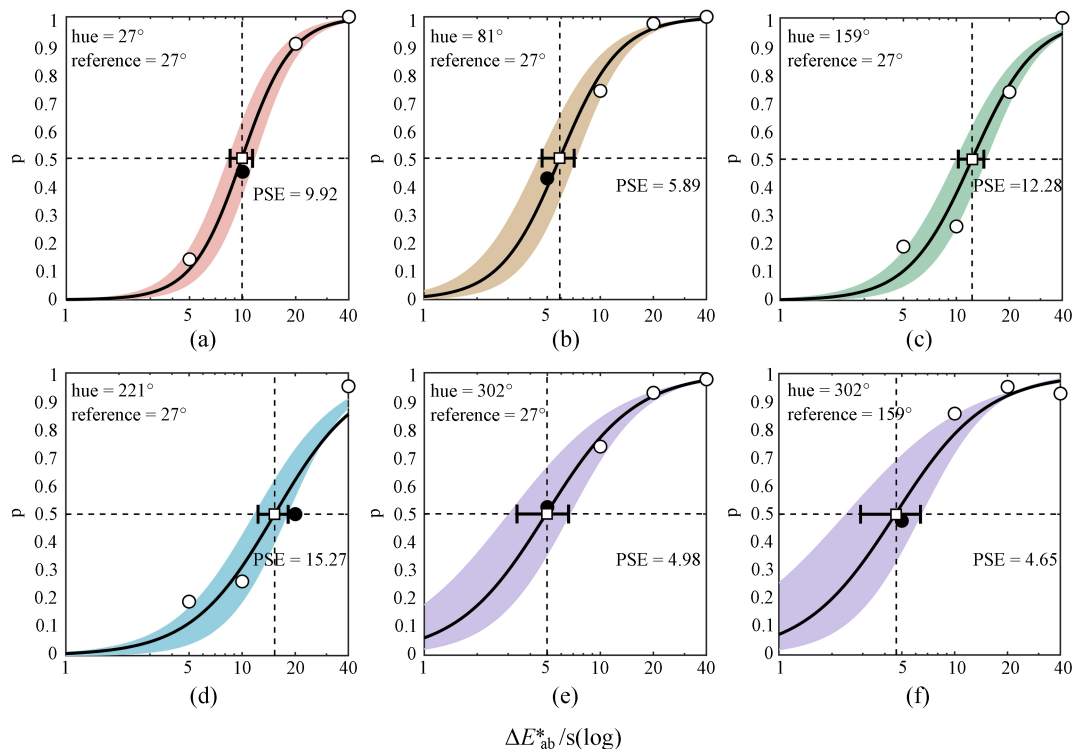


Figure 4.5: Scatter plot between the percentage (p) of the responses that ‘the comparison stimulus appeared faster than the reference stimulus’ and the speed of the comparison stimulus, for different base colors. The open circles represent the conditions where the judgements were significantly different from 50%; the filled circles represent the conditions where the judgements were not significantly different from 50%. The curves are the psychometric functions, as described in Equation 4.4, fitted based on the collected data. The shaded regions represent the possible psychometric functions that were derived using a non-parametric bootstrap method, with the error bars representing the 95% confidence interval for the estimated PSE value per base color. The different plots represent (a) base color of red; (b) base color of yellow; (c) base color of green; (d) base color of blue; (e) base color of purple; (f) base color of purple with the reference stimulus being switched to the base color of green.

A psychometric function, as described in Equation 4.4, was used to fit the data points of Figure 4.5 per base color:

$$f = \frac{1}{1 + e^{-(a+b \times \log_{10} x)}} \times 100\% \quad (4.4)$$

In this equation f represents the percentage of the judgments that “the comparison stimulus appeared faster than the reference stimulus,” x represents the speed of the comparison stimulus, and a and b are the fitting coefficients. The PSE value of each comparison stimulus is the speed at which the comparison stimulus appears to change at the same speed as the reference stimulus. Such a PSE value can be derived by solving the fitted psychometric function for f being 50%.

4

To estimate the variability of the two coefficients a and b of the psychometric function [Equation 4.4], we used a statistical technique based on Monte Carlo resampling, also known as a non-parametric bootstrap method [157] and used in Chapter 3. This technique assumed that for any given condition, the 42 responses made by the 21 observers reflected the underlying distribution of the entire population. Forty-two new responses were then randomly selected from this distribution, which was repeated for each of the four speeds until a new PSE was derived from this simulation-based procedure. This procedure was repeated 10,000 times, and the final PSE was the median of the 10,000 PSEs, while also the 95% confidence interval (CI) of the PSE can be calculated. The resulting six psychometric functions fitted using the data of the judgments made by the 21 observers, together with the PSE and the corresponding 95% CI as derived from the Monte Carlo resampling method, are shown in Figure 4.5.

For the conditions, in which the comparison stimulus had the base color of red, the derived PSE value was $9.92 \Delta E_{ab}^*/s$, as illustrated in Figure 4.5(a), which was not significantly different from the speed of the reference stimulus (i.e., $\Delta E_{ab}^*/s$). This result indicated a high reliability of the judgements made by the observers.

The PSE shown in Figs. 4.5(b)–4.5(e) indicates that the color transitions around the base color of yellow, green, blue, and purple needed to change their speed to 5.89, 12.28, 15.27, and $4.98 \Delta E_{ab}^*/s$, respectively, in order to appear as fast as the reference stimulus having the base color of red at a speed of $10 \Delta E_{ab}^*/s$. Moreover, comparing Figs. 4.5(e) and 4.5(f) suggests that the PSE of the comparison stimulus with the base color of purple is very similar for the reference stimulus with the base color of red (i.e., $4.98 \Delta E_{ab}^*/s$) as for the reference stimulus with the base color of green (i.e., $4.65 \Delta E_{ab}^*/s$).

4.4. DISCUSSION

To allow a direct comparison between this study and our study in Chapter 3, the PSEs and the 95% CIs derived from both studies are combined in Figure 4.6. It can be observed that the results of both experiments were highly comparable, with overlapping 95% CIs.

Based on the law of transitivity, which was also assumed in Chapter 3, the perceived speed of two stimuli (A and B) should be equal if another stimulus (C) is perceived to have the same speed as each of them (i.e., if $A = C$ and $B = C$ then $A = B$). As shown in the dashed box in Figure 4.6, the two reference stimuli with the base colors of red and green resulted in statistically the same PSEs for the comparison stimulus with the base color of purple, suggesting the stimuli with the base colors of red and green had similar perceived speed. However, their perceived speeds were statistically different when they were directly compared to each other, as illustrated by the open green bar in Figure 4.6, which was not

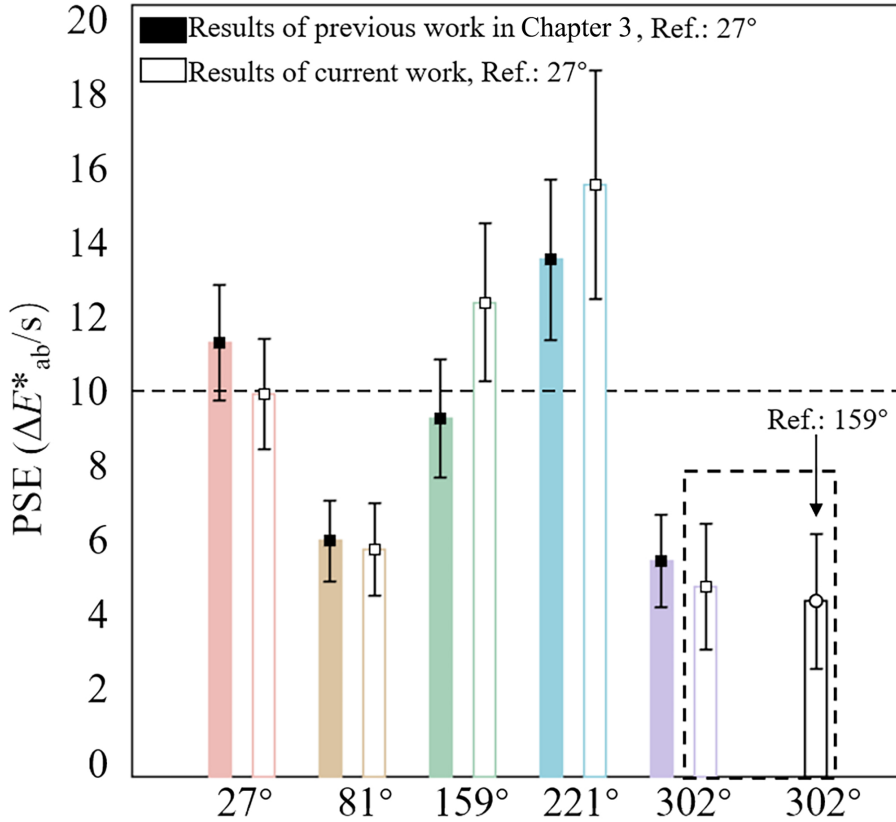


Figure 4.6: Comparison of the PSE values derived in this study (open bars) with those of our previous research [61] (solid bars). The two bars highlighted in the dashed box suggest that the PSE values for the base color purple derived for different reference stimuli are not significantly different.

consistent to our earlier findings in **Chapter 3**. As described above, this difference was likely due to using opponent colors without a complete adaptation. In **Chapter 3**, we used an adaptation interval of 10 s between the reference and comparison stimuli, while the interval was 0.5 s in our current experiment (see Figure 4.4). As reported in the literature, after viewing a colored stimulus, the perceived hue of the subsequent stimulus would be shifted in the opposite direction to the previous stimulus, which is caused by various physiological processes, such as neural adaptation and bleaching at photoreceptor level [3, 10]. This adaptation effect gradually disappears in 1–3 s [1], and thus, it is expected to have a larger impact in the current study than in our previous one. Since the adaptation effect interfered with the color transition, it is difficult to disentangle it from the appearance of the subsequent color transition.

Assuming the law of transitivity, it can be calculated that the color transition with the base color of purple at a speed of $4.65 \Delta E_{ab}^*/s$ is perceived as fast as the transition with the base color of green at a speed of $10 \Delta E_{ab}^*/s$ [see Figure 4.5(f)], which corresponded to a color transition with the base color of red at a speed of 8.14 [i.e., $10/(12.28/10)$] $\Delta E_{ab}^*/s$. Figure 4.5(e), however, showed that a color transition with the base color of purple at a speed of $4.98 \Delta E_{ab}^*/s$ was perceived as fast as a color transition with the base color of red at a speed of 9.34 [i.e., $10/(4.98/4.65)$] $\Delta E_{ab}^*/s$. Such a small discrepancy may be due to the accumulated error in our calculations, as also pointed out in **Chapter 3**. However, such an accumulated

error was acceptable, as the color transitions at the base color of red at speeds of $8.14 \Delta E_{ab}^*/s$ and $9.34 \Delta E_{ab}^*/s$ have overlapping 95% CIs [Figure 4.5(a)].

If CIELAB were temporally uniform, the ratio between $\Delta a^*/s$ and $\Delta b^*/s$ should be “1,” which means changes with the same Euclidean distance along the a^* and b^* directions should result in the same perceived speed of color change. Similar to our previous research in **Chapter 3**, we calculated the speed of change along the a^* and b^* directions (i.e., $\Delta a^*/s$ and $\Delta b^*/s$) of the stimuli at the speed of the PSE using a simple linear fit (see Figure 4.7), since $\Delta a^*/s$ and $\Delta b^*/s$ were not proportional to $\Delta E_{ab}^*/s$. We then optimized the scaling factor between $\Delta a^*/s$ and $\Delta b^*/s$, in order to minimize the deviations between the five PSEs. This was not only done in CIELAB, but also in the cone-excitation color spaces LMS and DKL [12], since these spaces were believed to better reflect the underlying mechanisms of human being’s temporal vision. The signals in the cone-excitation color spaces (i.e., $\Delta L/s$, $\Delta M/s$, $\Delta S/s$, $\Delta(L - M)/s$, and $\Delta[S - (L + M)]/s$) were calculated in a similar way, as described in **Chapter 3**. The L-, M-, and S-cone responses were calculated using the cone fundamentals reported in [133] and the measured spectral data at each step. The L-, M-, and S-cone values at each step were normalized by dividing them by the L-, M-, and S-values of the adaptation field (i.e., 4000 K), which was actually equivalent to multiplying them with a factor of 3.12, 3.85, and 11.73, respectively. These calculations were based on the assumption that the observers were completely adapted to the 4000 K illumination condition in such an experiment setup. However, in a past study [163], the slow adaptation time constants of neurons were measured for ΔS , $\Delta(L - M)$, and $\Delta(L + M + S)$ axes and they were all above 4 s (longer than our 3 s stimulus). This may suggest that the adaptation may not be complete with the stimuli being presented for 3 s. As pointed out in **Chapter 3**, there is a lack of a well-accepted model for temporal chromatic adaptation (which would change L, M, and S reference over time); the calculation of the cone responses to model speed perception needs to be further improved. In addition, as pointed out in [156], the desensitization for periodic color change along specific color-space directions may also influence the results. A better understanding of dynamic adaptation is worth exploring in future work.

It has to be noted that the optimization procedure adopted ten values in **Chapter 3** while it only adopted five values in the current work. The results of optimizing the scaling factors are shown in Table 4.2, and agree with **Chapter 3**. From these results we should conclude that the LMS space is not suitable for improving temporal uniformity as the M- and S-cone axes are simply omitted, which is not useful. For the optimization performed in the CIELAB and DKL space, the scaling factors—although the exact values are different from those calculated in **Chapter 3**—lead to the same interpretation that the b^* and $[S - (L + M)]$ axes contribute less to the perception of temporal modulations, compared to the a^* and $(L - M)$ axes, respectively. In **Chapter 3** and this study, 4000 K was adopted as the white point. More specifically, the 4000 K condition was used to continuously reset the observers’ state of adaptation in **Chapter 3**, while the 4000 K condition was employed as a static adaptation field in this study. In both cases, the L-, M-, and S-values of the stimuli were normalized by dividing them by the L-, M-, and S-values of the white point, which would influence the calculations of the weighting factors in the optimization routine. Similar experiments using different white points merit further investigations. In addition, adopting a different experimental paradigm, for example, keeping the stimulus constant while changing the speed of the illumination might be also interesting to explore.

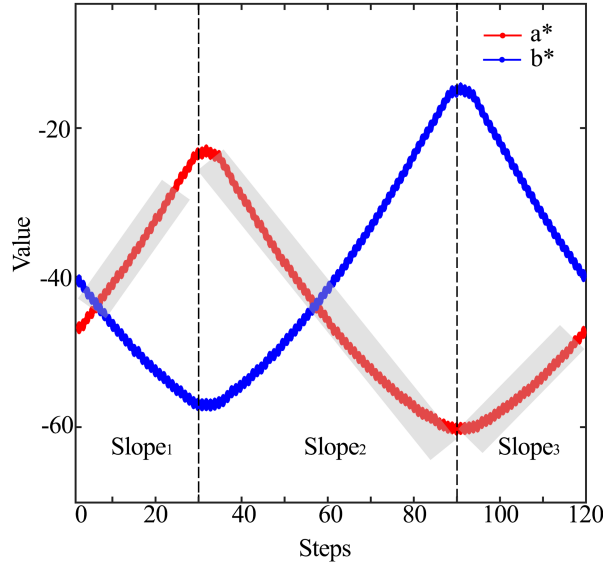


Figure 4.7: Illustration of the calculations of the change of speed along the a^* and b^* directions (i.e., $\Delta a^*/s$ and $\Delta b^*/s$) of the temporal color transition (L^* is constant). $Slope_1$, $Slope_2$, and $Slope_3$ covers the Steps 1–30, 31–90, and 91–120, respectively. The overall slope is given by $Slope = (Slope_1 - Slope_2 + Slope_3)/3$. The shaded areas in gray demonstrate the process of a linear fit to get one slope value expressed in unit of $\Delta a^*/s$. Figure retrieved from [61].

Table 4.2: Optimizations of color spaces to achieve a temporal uniformity

Color space	Inclusion of Scaling Factors	Optimization Results in Current Work	Optimization Results in Previous Work
CIELAB	$\Delta E_{ab}^*/s = \sqrt{(\Delta a^*/s)^2 + \alpha_{CIELAB}(\Delta b^*/s)^2}$	$\alpha_{CIELAB} = 0.04$	$\alpha_{CIELAB} = 0.404$
LMS	$\Delta LMS/s = \sqrt{(\Delta L/s)^2 + \alpha_{LMS}(\Delta M/s)^2 + \beta_{LMS}(\Delta S/s)^2}$	$\alpha_{LMS} = 0; \beta_{LMS} = 0$	$\alpha_{LMS} = 0; \beta_{LMS} = 0$
DKL	$\Delta DKL/s = \sqrt{[\Delta(L-M)/s]^2 + \alpha_{DKL}[\Delta(S-(L+M))/s]^2}$	$\alpha_{DKL} = 0.008$	$\alpha_{DKL} = 0.02$

4.5. CONCLUSIONS

In this study, 21 observers observed pairs of temporal color transitions around several base colors at several speeds, and indicated which of the two appeared to change faster. The responses allowed us to determine the point of subjective equality of speed between different color transitions. The stimuli were presented with a field of view of 4.3 deg with a static (adaptation) surround of 4000 K. Such a viewing condition is similar to what has been used to develop the CIELAB color space. We found that the color transitions around different hues needed to have different speeds, in terms of $\Delta E_{ab}^*/s$, in order to appear to change at the same speed. Specifically, the transition around the base color of yellow, green, blue, and purple needed to change at a speed of 5.89, 12.28, 15.27, and 4.98 $\Delta E_{ab}^*/s$, respectively, so that it appeared to change as fast as a transition around the base color of red at a speed of 10 $\Delta E_{ab}^*/s$. This finding indicates that $\Delta E_{ab}^*/s$ is not a suitable measure to predict the perceived speed of color changes, and that CIELAB is not a temporally uniform color space. We made preliminary attempts to derive scaling factors to improve the temporal uniformity of CIELAB and the cone-excitation spaces LMS and DKL. Given the current popularity of dynamic lighting, additional studies are needed to further improve the temporal uniformity of CIELAB.

5

An Experimental Comparison of Threshold Methods for Chromatic Flicker Detection: Accuracy, Precision and Efficiency

At present, all existing color spaces fail to accurately describe the perception of temporal color differences. Knowledge on the detection threshold of chromatic flicker under a wide range of conditions is essential for developing an accurate temporal color model. However, literature has shown inconclusive results on which psychometric method is most suitable for measuring detection thresholds. The aim of the present study was to compare the accuracy, precision and efficiency of five commonly used psychophysical methods for measuring the detection threshold of chromatic flicker. The methods employed were: (1) the classical 1-up-1-down staircase method with a yes-no task, (2) the weighted 3-up-1-down staircase method with a 2AFC task, (3) the method of constant stimuli with a 2AFC task, (4) the method of adjustment without reference, and, (5) the method of adjustment with reference. The chromatic flicker stimuli were temporally modulated with a square wave, i.e. two colors were shown alternately, where the mean chromaticity of the two colors was either 2700K, 4000K or 6500K, and the temporal frequency could be 2 Hz or 4 Hz. We found that each method has its advantages and disadvantages. Overall, the method of adjustment without a reference is recommended for future research on chromatic flicker detection, as it is very efficient, it is quite accurate compared to the assumed ground truth threshold and the variance of the threshold is comparable to those of the other methods.

5.1. INTRODUCTION

Nowadays, dynamic colored light is often being used in several applications, for instance, to create appealing atmospheres [40, 145, 97] or to enhance the immersive experience of displays, such as the Philips Amibilight TV [115]. Applications of dynamic light have been limited in the past, mainly because traditional lighting technologies did not allow easy control of colored light. The introduction of light emitting diodes (LEDs) in lighting has changed the playing field, since it enables inexpensive ways to easily create colored and dynamic light effects. These colored, dynamic light effects should have certain properties in order to be attractive to observers: the light should change smoothly over time, both in luminance and chromaticity, and at the desired (perceived) rate of change [121, 97, 120]. However, current knowledge on human perception of dynamic colored light is still insufficient to provide guidelines for pleasant implementations. Information that is missing concerns, among others, the perceived amount of color difference and the perceived rate of change of transitions from one color to another color. In other words, what is finally needed is a perceptually uniform color space for dynamic light. Chromatic flicker is a useful paradigm to study the sensitivity of human observers to temporal color differences.

5

Flicker is defined as “the impression of unsteadiness of visual perception induced by a light stimulus whose luminance or spectral distribution fluctuates with time...” [48]. The study of chromatic flicker dates back to the 1910s, when Ives and Troland used heterochromatic flicker photometry to determine when two colors match in brightness. Heterochromatic flicker photometry is based on the observation that two alternating lights of different color appear steady at frequencies above the chromatic critical fusion frequency (CCFF). The CCFF has been found to depend on the color pair [140] and the retinal illuminance [44, 140]. At frequencies below the CCFF, the sensitivity to chromatic flicker depends on the frequency, which can be described by the temporal contrast sensitivity function (TCSF). However, literature shows inclusive results on the shape of the TCSF. Some studies have found a bandpass function that peaks around 4 Hz [36, 122], whereas other studies have reported low pass characteristics [45, 34, 136]. The sensitivity has also been found to depend on the mean (adapting) luminance and stimulus size [136]. However, due to limited capabilities of previous lighting technologies, most studies have investigated mainly red-green chromatic flicker. Currently, LEDs enable us to create and study a large variation in chromatic flicker stimuli, which can therefore significantly enlarge our knowledge on the perception of dynamic colored light.

The perception of chromatic flicker is usually measured with a psychophysical method that measures the statistical estimate of the threshold above which the temporal fluctuation can be detected at a specific probability. Several methods for measuring thresholds have been used in literature. As pointed out by Treutwein [139], psychophysical methods should be evaluated in terms of costs and benefits according to the following three aspects:

- Accuracy, bias, or systematic error: these characteristics refer to the deviation of the measured threshold from the veridical or “true” threshold. In any real psychophysical experiment, the value of the true threshold is unknown, which makes it impossible to evaluate the bias of a particular psychophysical method. The bias of a method can only be determined in simulations.
- Precision: this is related to the variation of the threshold when it is measured repeatedly. A high precision corresponds to a small variation.

- Efficiency: this is related to the number of trials required to achieve a certain precision. A high efficiency corresponds to a low number of trials. The aim of our study is to compare the precision and efficiency of several psychophysical methods for measuring the detection threshold of chromatic flicker. In a later study, the most promising method will be used to investigate the effect of several parameters on the detection of chromatic flicker.

5.2. PSYCHOPHYSICAL METHODS

5.2.1. CLASSICAL PSYCHOPHYSICAL METHODS

Psychophysical experiments can be used to measure thresholds, which refers to the stimulus intensity that corresponds to a certain level of performance (e.g. 50% correct). Two types of thresholds can be distinguished: detection thresholds and discrimination thresholds [104]. Detection thresholds, also commonly called absolute thresholds, are defined as the minimal intensity of a stimulus that an observer can just detect. Discrimination thresholds, also called difference thresholds, refer to the minimal change in intensity that an observer can just detect. Obviously, in a discrimination experiment at least two alternative stimuli have to be presented for a given trial, one of which being the reference; the observer is asked to distinguish between the reference and the other stimulus. In a detection experiment only one stimulus has to be presented and the observer is asked to evaluate this stimulus. However, any number of alternative stimuli can be presented, for instance, including a blank stimulus. In that case, the observer is asked to distinguish between the blank and the other stimulus.

Fechner [24] proposed three types of methods for estimating perceptual thresholds, now called *classical* psychophysical methods: the method of constant stimuli, the method of limits and the method of adjustment. The methods differ in the decision task of the observer and in the stimulus presentation mode. The first two methods employ a judgement task, where the experimenter controls the presentation of the stimuli using a specific procedure and the observer makes a judgment. The method of adjustment uses an adjustment task, where the observer adjusts the stimuli to satisfy a perceptual criterion specified by the experimenter (e.g. to find the just noticeable stimulus intensity). Adjustment experiments are usually more efficient than judgement experiments because they directly yield the physical stimulus that corresponds to the criterion. However, observers may interpret the perceptual criterion differently, which may confound the results. Judgement tasks measure human response as a function of an experimental parameter, e.g. the intensity of a stimulus, to find the value that corresponds to a perceptual criterion. Judgement experiments can differ in the type and number of alternative responses that the observer is allowed to give. In a yes-no task, sometimes called simple forced choice task, the observer has two response options: “yes, the stimulus or phenomenon was present” or “no, the stimulus or phenomenon was not present”. In a two-alternative forced choice task (2AFC) two stimuli are presented, one of which is the target stimulus, and the observer is forced to choose which of the stimuli is the target one. Similarly, also n-alternative forced choice experiments exist.

METHOD OF CONSTANT STIMULI

In the method of constant stimuli, a number of stimulus intensities, which are likely to include the threshold value, are chosen by the experimenter. Each stimulus intensity is

presented multiple times (not less than 20, as suggested by Ehrenstein and Ehrenstein [18]; at least 10, as suggested by Fairchild [21]). In practice, some “catch trials” are included where there is no stimulus presented (or a stimulus with zero intensity) to check if participants understand the task and respond as expected. The resulting set of stimuli is presented one by one to an observer in a random order. After each stimulus presentation, a response is recorded (e.g. “yes” or “no”). The frequency or percentage of one response category plotted against the stimulus intensity results in a psychometric function (see Figure 5.1). A psychometric function usually is a sigmoid curve and not a step function, corresponding to the fact that perception is a statistical property. The internal representation of a stimulus intensity is usually described by a variable with a normal distribution, representing internal noise [18]. In a yes-no task, the threshold is usually defined as the stimulus intensity that corresponds to 50% “yes” responses, which is equivalent to 75% correct responses in a 2AFC task.

5

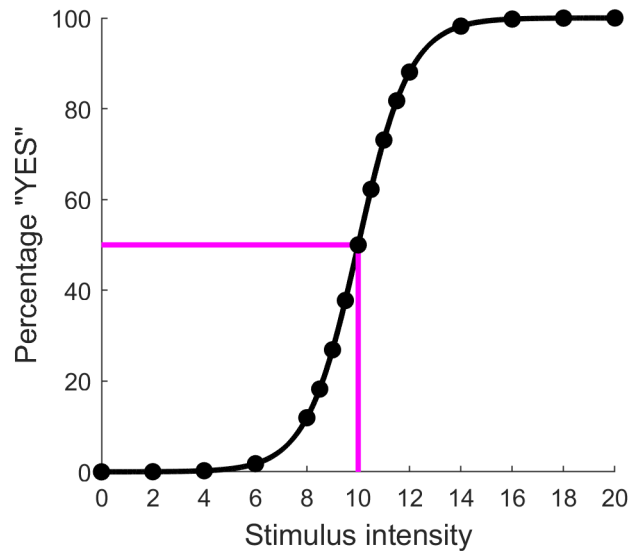


Figure 5.1: Illustration of a psychometric function: the percentage of responses “yes” is plotted against stimulus intensity.

The advantage of the method of constant stimuli is that it provides an estimate of both the threshold and the variance of the distribution of the internal stimulus representation, given by the slope of the psychometric function. Moreover, it is often assumed to be the most accurate method for measuring thresholds. However, Levison and Restle [76] showed that the psychometric function is shifted toward the center of the chosen intensity levels, because of the tendency of participants to use available responses with equal frequency [119]. Intensity levels are usually chosen based on reported thresholds from literature or pilot studies. Hence, results can be influenced by pre-experimental expectations. Researchers also have different opinions about the efficiency of the method of constant stimuli. Based on Monte Carlo simulations, Simpson [126] claimed that the method of constant stimuli is more efficient than the method of limits for 100 trials or less. Watson and Fitzhugh [152] indicated some unreasonable assumptions in Simpon’s simulation model and drew the opposite conclusion, namely that the method of constant stimuli is inefficient.

METHOD OF LIMITS

In the method of limits, stimuli are presented in either ascending or descending order until the observer indicates that a predefined criterion is met. In the context of measuring detection thresholds, for ascending series, the first stimulus is presented at an imperceptible intensity and the intensity gradually increases until the observer can perceive the stimulus. For descending series, a clearly perceptible stimulus is presented initially and the intensity gradually decreases until the observer cannot perceive the stimulus any longer. The same procedure is used for measuring discrimination thresholds, but then the intensity of the reference stimulus is kept constant, while the intensity of the target stimulus decreases or increases.

This method is easy to implement, however, the observer may get used to reporting a response and may continue reporting that same response even beyond the threshold, which is referred to as the error of habituation. Conversely, the observer may anticipate that the threshold stimulus is about to come and may make a premature judgement, which is referred to as the error of anticipation [76]. As a result, ascending and descending series often yield small, yet systematic differences in thresholds [18]. A possible solution to overcome this is presenting both ascending and descending series, and averaging the corresponding thresholds [18, 25]. Another solution is to use adaptive methods, as described in Section 5.2.3.

METHOD OF ADJUSTMENT

In the method of adjustment, the observer controls the stimulus intensity towards a level that meets a criterion specified by the experimenter. The method is widely used, but may result, as mentioned before, in a large variance because observers can interpret the criterion differently. The latter issue can be alleviated by training observers to raise or lower their criterion [104]. Similar as with the method of limits, the error of anticipation may impact thresholds measured with the method of adjustment. Thus the intensity of the first stimulus may influence the measured threshold [50]. In addition, limited efforts have been done to evaluate the influence of the stimulus intensity that is being presented first. A big advantage of the method of adjustment is that it can be quite efficient.

All psychophysical methods described above suffer from one or more deficits [23]. The methods with either a yes-no task or an adjustment task have no control over the decision criterion of the participant, which means that the sensitivity of the sensory process is confounded with the response criterion [41]. The alternative is to use a 2AFC task or to apply Signal Detection Theory [100]. In the latter case, both trials with a signal and trials without a signal are presented. Based on the proportions of the Hit rate (i.e. the signal is present and the observer responds “yes”) and the False Alarm rate (i.e. the signal is absent and the observer responds “yes”), the sensitivity (called d' -prime) and the decision criterion (called β) can be estimated. However, this means that the number of stimuli may be doubled. Another deficit of the classical methods is that a large amount of stimuli are presented far from threshold, leading to a waste of data [23]. To remedy this deficit, adaptive psychophysical methods have been proposed.

5.2.2. ADAPTIVE PSYCHOPHYSICAL METHODS

In adaptive psychophysical methods, the intensity of the stimulus that is presented is not predefined but depends on the response(s) of the observer to a previous trial(s). Adaptive

algorithms for the measurement of sensory functions have been introduced in the 1930s [75], and various variations of the method have been used since then. Treutwein [139] presents a comprehensive overview of adaptive methods including criteria for choosing among the various techniques. Most methods yield only a threshold value, and in some cases, other characteristics of the psychometric function (such as the slope) can be estimated as well. Leek [75] has reviewed the utility of the adaptive methods and the circumstances under which one method might be more appropriate than another method, based on both simulations and perception experiments. What all methods have in common is that they need to address the following issues [62]:

- the starting stimulus value (e.g. amplitude or intensity);
- a stimulus placement rule for the next trial, including the step size (i.e. the difference in stimulus value) and the tracking algorithm (i.e. when the stimulus should change);
- a termination rule;
- a rule to estimate the psychometric parameters (e.g. threshold and slope) based on previous responses.

The review by Leek [75] classifies the adaptive methods into three categories: parameter estimation by sequential testing (PEST), maximum-likelihood adaptive methods and staircase methods. Since we use the staircase method in our experiment, we describe this method more explicitly in the next sub-section. More information on the PEST and maximum-likelihood method can be found in Taylor and Creelman [137] and Watson and Pelli [153].

STAIRCASE METHOD

The staircase method, also known as the “up-down” method, was developed by Dixon and Mood [14]. It is a variation of the method of limits, where the stimulus level does not continuously decrease or increase, but either decreases or increases depending on whether the phenomenon is observed in the previous trial(s) or not. In the simple up-down method (1-up-1-down) the stimulus intensity is increased after each negative response and decreased after each positive response. This converges to the 50% performance point of the psychometric function. Levitt [77] proposed the transformed up-down method, where the stimulus level changes depending on the responses of two or more of the preceding trials, making it possible to measure other points of the psychometric function. For example, a transformed n -up- m -down method means that the stimulus level increases with a step size D after each n continuous negative (or incorrect) responses and decreases with a step size D after each m continuous positive (or correct) responses. Kaernbach [51] described a weighted up-down method that could be used to estimate even more points on the psychometric curve, by removing the requirement that the step size D^- for decreasing the stimulus equals the step size D^+ for increasing the stimulus. Note that a weighted n -up- m -down method actually corresponds to a 1-up-1-down method, but with $D^+/D^- = n/m$.

Also, other variables of the staircase method, such as the starting value, termination criterion, the number of reversal points (i.e., the number of times in which the stimulus intensity changes from increasing to decreasing, or vice versa) used to estimate the threshold and the number of staircases run to obtain a final estimate have been widely discussed. For

example, García-Pérez [26, 27, 31, 28, 30] extensively studied both yes-no staircases and forced-choice staircases using simulation techniques, and demonstrated that using steps up and down of different size and in a certain ratio stabilizes the performance of a staircase.

5.2.3. COMPARISON OF THRESHOLD METHODS

Many studies have been conducted to compare different methods for measuring thresholds and different configurations of these methods (such as the starting value). Since simulations give fast results at a low cost, and are the only way to evaluate the accuracy of a psychometric method, most of these studies are based on computer simulations. In simulations, an ideal observer is assumed, while in reality, many factors (such as the instructions of the experimenter) may influence the observer's behavior. Besides, an observer may make decisions in irrational ways [41]. All these uncontrollable factors may lead to differences between simulations and real experimental results. Therefore, experiments with real observers are needed to confirm the conclusions of simulations.

There are a couple of studies in literature that have used real observers for comparing different threshold methods. For example, Wier, Jesteadt, and Green [158] compared the method of adjustment, the transformed 1-up-2-down staircase method and the method of constant stimuli to measure auditory frequency discrimination. Estimates of the threshold using the method of adjustment were on average a factor 2 smaller than those of the other two methods. They argued that the adjustment method depended on a number of factors (such as the step size to adjust the stimulus levels and the subjects' strategy for terminating the trial) that influenced the results. Podlesek and Komidar [106] compared the same three methods for measuring the displacement of a target moving in a frontal plane. The displacement obtained with the method of adjustment was larger than the displacement obtained with the other two methods. Also, the variability was largest for the method of adjustment. This shows that the method of adjustment should be used with caution. Rammsayer [107] performed an experiment comparing the weighted 3-up-1-down method with the transformed 1-up-2-down method for an auditory temporal discrimination task. He reported that the weighted up-down method was more efficient than the transformed up-down method as long as less than 40 trials were presented. In a recent paper, Eisen-Enosh et al. [19] compared the method of limits, the method of constant stimuli and the weighted 3-up-1-down staircase method to measure the critical flicker-fusion frequency. Their results demonstrate that all three methods are able to reliably measure the critical flicker frequency with high repeatability, but that the threshold agreement is highest between the staircase and method of constant stimuli, while the latter had the lowest efficiency.

5.2.4. OBJECTIVES AND HYPOTHESES

The aim of our study is to compare the precision and efficiency of five commonly used psychophysical methods for measuring the detection threshold of chromatic flicker. The above-mentioned literature demonstrates that there is no single most appropriate method for estimating thresholds. The most appropriate method depends on many (uncontrollable) factors. Therefore, we directly compared:

1. Method 1: the classical 1-up-1-down staircase method with a yes-no task
2. Method 2: the weighted 3-up-1-down staircase method with a 2AFC task

3. Method 3: the method of constant stimuli with a 2AFC task
4. Method 4: the method of adjustment without a reference
5. Method 5: the method of adjustment with a reference

In methods 1 and 4, only one chromatic flicker stimulus was presented at a time, and therefore these methods directly measure the detection threshold. In the other methods, two stimuli were presented next to each other, i.e., a reference stimulus without flicker and a test stimulus with chromatic flicker. These methods measure the discrimination threshold with respect to a static stimulus, which in fact corresponds to the detection threshold of chromatic flicker. The method of constant stimuli was chosen in order to determine the entire psychometric function underlying chromatic flicker visibility. The staircase methods were chosen because they are often found to be relatively efficient, whereas the precision depends on the specific task. The method of adjustment was included because it is usually fast, but at the expense of precision and/or accuracy.

5

As mentioned above, participants have their own criterion to decide when a stimulus is visible or not. Some participants may be very conservative and give a positive response only when they are very confident, whereas other participants might be more liberal. Methods 2 and 3 use a 2AFC procedure which reduces the effect of the decision criterion, since participants are able to compare the stimuli and make a judgment on the difference. As these two methods appear the same from the participants' point of view, they are expected to yield the same estimate of the threshold, at least when the staircase method truly converges to the 75% performance point on the psychometric function. Methods 1, 4 and 5 do not control for participants' criterion and, therefore, the thresholds are expected to deviate from those measured with Methods 2 and 3. It is not clear from literature whether the threshold is expected to increase or decrease, since the conclusion of Wier, Jesteadt, and Green [158] contradicts the conclusion of Podlesek and Komidar [106].

Apart from affecting the threshold, the different methods may also affect the variance of the threshold. Methods 2 and 3 are expected to have the smallest variance since the variance of these methods is mainly determined by the difference in sensitivity between participants and not by the difference in decision criterion, as is the case for the other methods. In Methods 4 and 5 participants have full control over the stimuli themselves, which is expected to result in an additional source of variance. The variance of Method 5 is expected to be slightly smaller than the variance of Method 4, because of the presence of the reference.

Finally, the various methods are also expected to differ in efficiency. Methods 4 and 5 are known to be usually very efficient, since participants determine themselves when they reach the threshold. Method 5 is expected to be somewhat less efficient than Method 4, since it takes some time to compare the test stimulus with the reference. Method 3 is expected to be the least efficient one, since it shows equal amounts of stimuli at places nearby and far away from the threshold.

So, to summarize we formulate the following hypotheses:

- The mean of the threshold: Method 2 = Methods 3; Method 1 = Method 4 = Method 5
- The variance of the threshold: Method 3 = Method 2 < Method 1 < Method 5 < Method 4

- The efficiency of the method: Method 4 > Method 5 > Method 1 > Method 2 > Method 3

5.3. METHODS

5.3.1. EXPERIMENTAL DESIGN

In this experiment, the detection threshold of chromatic flicker was measured for light that was temporally modulated around a base color at a certain frequency. Five different methods were used to measure the threshold for three base colors (2700K, 4000K and 6500K) and two temporal frequencies (2 Hz and 4 Hz). Each condition was repeated several times, depending on the methodology, as described below. The experiment had a full factorial within-subject design.

5.3.2. EXPERIMENTAL SETUP

The light source used in the experiment was an LED luminaire containing 3 red, 3 green and 3 blue LEDs behind a diffuser. The light was controlled by a customized IP-addressable circuit board connected to a computer through an Ethernet cable. The LEDs were driven by means of pulse width modulation (PWM) at a driving frequency of 500 Hz and 11-bit levels. The driver accepted RGB values, ranging from 0 to 2047, in the device dependent color space of the LEDs. The target stimuli were defined in CIE 1976 UCS (u' , v') and transformed via XYZ to the RGB values of the LEDs.

The luminaire was placed in a box (height: 1.5 m, depth: 0.8 m, width: 1.5 m) with a circular opening of 26.4 cm diagonal in order to make the light stimuli more uniform (see Figure 5.2). The inner surfaces of the box were smooth and colored natural white. Participants looked at the light stimuli from a predefined position of 1.5 m from the front of the box, which resulted in a visual angle of 10-degrees. Most existing color models are based on either a 2-degrees or 10-degrees field of view. As our focus is on realistic lighting applications, we chose a visual angle of 10-degrees. The inner edges of the box and the luminaire itself were not visible. The stimuli were presented at a luminance of 6.9 cd/m². The luminance of the stimuli varied slightly within the circular opening with a maximum variation of 2.2 cd/m².

Two identical boxes and LED luminaires were made to enable the use of a two-alternative forced choice (2AFC) procedure. The visual angle between the two boxes was approximately 45 degrees horizontally (see Figure 5.2(b)). In order to avoid head movements of the participants, a chinrest was used. Additionally, a standard keyboard was provided as input device for the participants.

5.3.3. STIMULI

The light of the stimuli was temporally modulated with a square wave, which means that two colors were shown alternately with a certain time interval. The chromaticity of the two colors varied around a base color in a predefined direction. Three base colors were selected on the black body curve (BBC), characterized by their color temperature of 2700K, 4000K, and 6500K. These values were chosen as they are common in lighting applications. The corresponding chromaticity coordinates of the base colors in the CIE 1976 UCS color space, computed with the equations of Krystek [66], are presented in Table 5.1. Since the main purpose of this study was to investigate the effect of methodology, only one direction of the

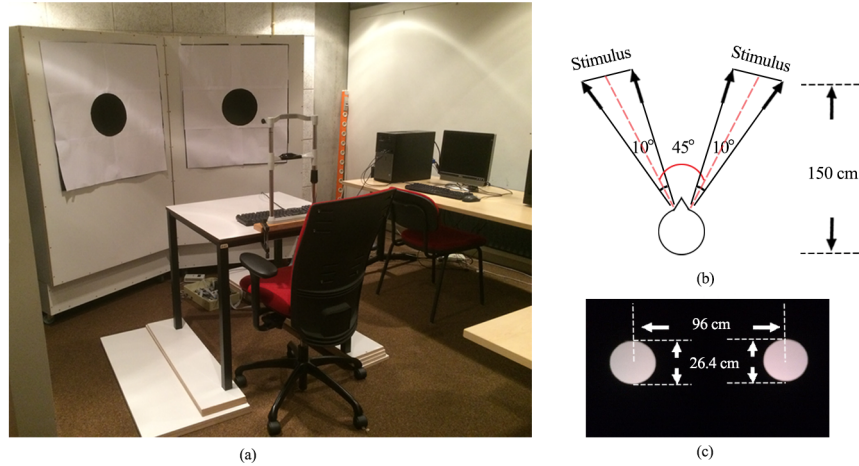


Figure 5.2: (a) Overview of the experimental setup. (b) Top view of the participant and stimuli. (c) Front view of the stimuli.

5

chromaticity modulation was investigated, namely the direction perpendicular to the BBC in the CIE 1976 UCS color space (see Figure 5.3). The chromaticity coordinates of the two alternating colors depended on the modulation amplitude.

Table 5.1: The chromaticity coordinates of the three base colors and their modulation direction in CIE 1976 UCS color space and the two modulation frequencies

Base color	u'	v'	Modulation Direction	Frequency
$C_1 = 2700K$	0.262	0.528	$\theta_1 = 113^\circ$	2 Hz, 4 Hz
$C_2 = 4000K$	0.225	0.502	$\theta_1 = 136^\circ$	2 Hz, 4 Hz
$C_3 = 6500K$	0.200	0.465	$\theta_1 = 153^\circ$	2 Hz, 4 Hz

Even though the physical luminance of the two alternating colors was equal, participants could perceive the luminance to be different due to individual deviations from the luminous efficiency function of the standard observer [82, 113]. Therefore, it might be possible that participants detect luminance flicker rather than chromatic flicker. Since people are more sensitive to chromatic flicker than luminance flicker at low frequencies [136], we choose modulation frequencies of 2 Hz and 4 Hz in order to minimize a possible influence of perceived luminance flicker.

5.3.4. METHODOLOGY

The following methods were used to measure the detection threshold of chromatic flicker: (1) the classical staircase method with a yes-no task, (2) the weighted 3-up-1-down staircase method with a 2AFC task, (3) the constant stimuli method with a 2AFC task, (4) the method of adjustment without a reference, and (5) the method of adjustment with a reference. Parameters related to the specific implementation of these methods (for example, the number of repetitions and the starting point) were determined based on multiple pilot experiments. Design choices were made with additional criteria in mind, such as taking

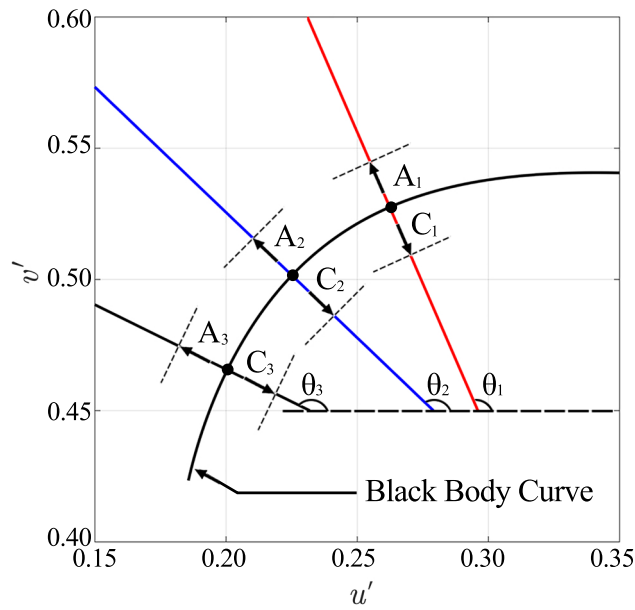


Figure 5.3: Representation of the three base colors (C_1, C_2, C_3), an example modulation amplitude (A_1, A_2, A_3), and the modulation directions ($\theta_1, \theta_2, \theta_3$) together with the black body curve in the CIE 1976 UCS color space.

care that comparable number of stimuli were used for all methods and that all experimental sessions required a comparable amount of time of about 30 minutes (in order to avoid an effect of experimental fatigue between the different methods).

METHOD 1: CLASSICAL STAIRCASE WITH YES-NO TASK

The general procedure of this method is described in Section 5.2.2. The first stimulus that was presented had a clearly visible chromatic modulation with an amplitude of $0.05 \Delta(u', v')$. The stimulus presented in the n -th trial had an amplitude indicated by A_n . If the response to this stimulus was “no”, the amplitude A_{n+1} of the stimulus at the $(n + 1)$ -th trial was increased by a factor F_+ , and if the response was “yes”, the amplitude was decreased by a factor F_- . The actual factor depended on the passed number of reversal points as indicated in Table 5.2; at the start of the staircase the factor was 1.1^5 (or 1.1^{-5}) and after 4 reversal points it was reduced to 1.1^2 (or 1.1^{-2}). There were two reasons for changing the modulation amplitude by means of a factor instead of a fixed step; the first being that negative amplitude values could be avoided, the second being that it was more logical to change the amplitude on a logarithmic scale, according to the Weber–Fechner Law. After nine reversal points the staircase procedure ended, and the threshold was calculated as the average over the last four reversal points. For each base color, six staircases (i.e. 2 frequencies \times 3 repetitions) were intermingled. Therefore, participants had no idea which staircase they were executing, and they could not anticipate on their responses.

METHOD 2: WEIGHTED STAIRCASE WITH 2AFC TASK

As described in Section 5.2.2, this method measures the 75% discrimination threshold of a chromatic flicker stimulus with a static reference stimulus, and as such the threshold can be interpreted as the detection threshold of chromatic flicker. The procedure was very similar

Table 5.2: Parameters for the implementation of the staircases for Method 1 and Method 2. Note that the value of α is changing depending on the number of reversal points. From 1 to 4 reversal points α is 5, while from 5 to 9 reversal points α is 2.

Method	Up/down rule	F+	F-	Starting value
1	1/1	$\frac{A_{n+1}}{A_n} = 1.1^\alpha$	$\frac{A_{n+1}}{A_n} = 1.1^{-\alpha}$	$\Delta(u', v') = 0.05$
2	3/1	$\frac{A_{n+1}}{A_n} = 1.1^{3\alpha}$		

to that of Method 1. The staircase started with a clearly visible chromatic modulation of $0.05 \Delta(u', v')$. Also, the amplitude was increased after a “no” response and decreased after a “yes” response. However, the factor $F+$ was different, as shown in Table 5.2.

METHOD 3: CONSTANT STIMULI WITH 2AFC TASK

This method aims to measure the entire psychometric curve and uses a predefined set of amplitudes. In our experiment, nine modulation amplitudes $A_{cs} = \{A_i | i = 1, 2, 3, \dots, 9\}$ were selected. The first amplitude corresponded to $A_1 = 0$, which is a static stimulus, and the last amplitude corresponded to $A_9 = 0.05 \Delta(u', v')$, which corresponds to a stimulus with obvious flicker. These stimuli served as catch trials (as explained in Section 5.2.1) to check if people understood the task. The other seven amplitudes were different per participant and were calculated based on the data obtained from both staircase methods, i.e., Method 1 and Method 2. For each participant, a psychometric function was fitted (see Section 5.3.7) on the combined data of the two methods using a generalized linear model (GLM). A_2 and A_8 were deduced from this fit and corresponded to the amplitude for which the percentage “yes” responses was 15% and 85%, respectively. The remaining amplitudes from A_3 to A_7 were linearly interpolated. All amplitudes were presented ten times. In order to reduce order effects, as found in literature [87], the resulting 180 stimuli of a base color were grouped in 10 blocks; within each block the nine amplitudes for the two frequencies were presented in a random order.

METHOD 4: METHOD OF ADJUSTMENT WITHOUT REFERENCE

In this method, participants were instructed to increase or decrease the amplitude of the chromatic flicker stimulus and to select the smallest amplitude at which the flicker was just visible. The method started either from a clearly visible flickering stimulus with a chromatic modulation of $0.05 \Delta(u', v')$, or from a visually non-flickering stimulus with an amplitude of $0.0004 \Delta(u', v')$. The modulation amplitude could be changed with a large or smaller factor, depending on how close the participants considered themselves to be near the threshold. The large factor was:

$$\frac{A_{n+1}}{A_n} = 1.1^{\pm 5}$$

whereas the small factor was:

$$\frac{A_{n+1}}{A_n} = 1.1^{\pm 2}$$

where in both cases A_n refers to the amplitude of the n -th trial; + and - correspond to increasing and decreasing the amplitude respectively. Each participant performed the adjustment three times starting from an amplitude of $0.05 \Delta(u', v')$, and three times starting from an amplitude of $0.0004 \Delta(u', v')$. So, they performed six adjustments per base color and frequency.

METHOD 5: METHOD OF ADJUSTMENT WITH REFERENCE

This method only differed from Method 4 in the sense that two stimuli were presented: a static reference stimulus and a chromatic flicker stimulus. Participants were instructed to increase or decrease the amplitude of the chromatic flicker stimulus with either a large or a smaller factor until they reached the smallest amplitude at which the difference between the two stimuli was just visible. Again, all participants performed three adjustments starting from a clearly visible flickering stimulus with an amplitude of $0.05 \Delta(u', v')$, and three adjustments from a visually non-flickering stimulus with an amplitude of $\Delta(u', v')$.

5.3.5. PARTICIPANTS

Thirty-two young adults (26 males and 6 females) participated in the experiment. Their age ranged from 19 to 29 years ($M = 23$, $SD = 2.5$). Among the participants, twenty-one persons executed all five methods (of which 19 were male and 2 were female, aging from 19 to 29 years, $M = 23$, $SD = 2.4$). The other eleven persons only participated in Method 1, 2 and 3. Participants that were susceptible to migraine and/or epileptic seizures were excluded from participation. All participants had normal color vision, as measured with the Ishihara test for color deficiency. Six of the participants had previous experience with visual perception experiments.

5.3.6. PROCEDURE

The participants executed the methods in separate sessions on different days, taking into consideration that measuring all methods at once would be too tiring. All participants started with Method 1, followed by Method 2 and then Method 3. Ten of the participants continued with Method 4 and then Method 5, and, eleven participants continued with Method 5 and then Method 4.

Before the start of the first session, participants were asked to sign an informed consent form, in which they were informed that they could quit the experiment anytime. Another form was used to collect information on: demographics (age, gender), whether they had glasses or lenses and whether they had experience with visual perception experiments before. Then they were asked to perform an Ishihara color blindness test and a visual acuity test. In each session, the task of the participant was explained both in text and orally by the experimenter. Then the participants were instructed to sit down in the chair at the predefined position. They could adjust the chinrest and the height of the chair to a comfortable position before the experiment started. They were instructed to look at a fixation point in the center of the circular opening when judging the stimulus (both when one and two boxes were used). After the lights in the room were turned off, some demonstration stimuli were shown to the participants and they could practice for a few trials. Whenever necessary, the experimenter answered questions asked by the participants before the actual start of the experimental trials.

In Method 1, one stimulus was presented at a time in the box at the left side. Participants were instructed to use the left and right arrow buttons to indicate if the stimulus was flickering or not. They were asked to respond quickly (i.e. based on their first impression) but accurately. The staircases of the three base colors were presented in three consecutive blocks, as indicated in Figure 5.4. The order of the base colors was counterbalanced between participants to avoid a possible confounding factor. For each block, first the corresponding base color was shown for one second, and then a beep sound indicated the presentation

of the chromatic flicker stimulus. Participants could look at the stimulus without any time restriction. After pressing one of the two buttons, the same base color was shown again for one second, followed by the next chromatic flicker stimulus, as determined by the staircase procedure.

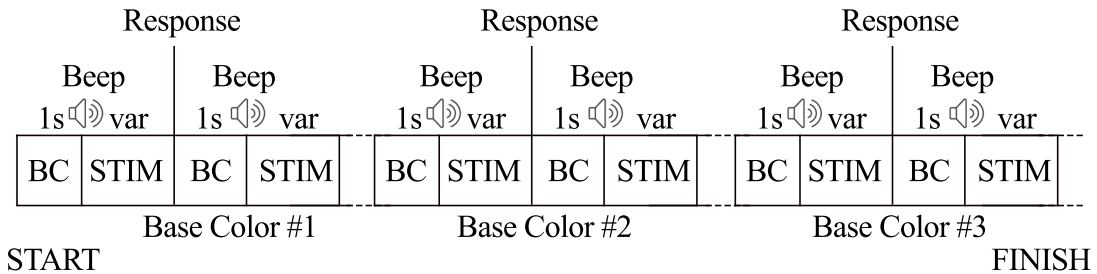


Figure 5.4: Time line for the psychophysical procedure of Method 1, Method 2, and Method 3; BC stands for base color, while STIM stands for stimulus (which can be with or without a reference.)

5

In Method 2 and Method 3, the flicker stimulus and the reference were presented simultaneously in the two neighboring boxes. Again, the three base colors were assessed one after the other (in a counterbalancing order) with the same time line as shown in Figure 5.4. Participants were instructed to use the left and right arrow buttons to indicate whether the flicker stimulus was shown on the left side or the right side, while for each trial the flicker stimulus was presented randomly at one of the sides.

In Method 4, one stimulus was presented at a time in the box at the left side. Again, the adjustments for the three base colors were executed one after the other in a counterbalancing order. Each adjustment started with presenting the corresponding base color for 10 seconds, after which a beep sound indicated that the chromatic flicker stimulus with either a very high or a very low amplitude was ready to be adjusted, as shown in Figure 5.5. Participants could use the up and down arrow buttons to increase or decrease the amplitude of the flicker stimulus with a large step, while the left arrow and right arrow buttons could be used to change the amplitude with a small step. Participants were instructed to find the stimulus, for which they just saw flicker. When they reached the system's lower boundary (at $0.00004 \Delta(u', v')$) or higher boundary (at $0.05 \Delta(u', v')$), they heard a warning signal to indicate that they could not go lower or higher, respectively. Participants were free to use their strategy but they were suggested to look at the final stimulus for one or two seconds before deciding to press the Enter button to end the adjustment procedure. A beep sound indicated the start of a new adjustment. All 12 conditions per base color (i.e., 2 frequencies \times 2 starting amplitudes \times 3 repetitions) were presented in a random order.

In Method 5, a flicker stimulus was presented in the box at the left and a static stimulus (the reference) in the box at the right side. Participants were explicitly told to compare the two stimuli in their assessment of seeing flicker. The rest of the procedure was equal to that of Method 4.

After each method (except for Method 3, because the task of the participants was the same as for Method 2), participants were asked to answer some questions regarding their experience with the method. During the experiment, we stored all threshold data, but also the amplitude of all stimuli that were presented together with the time log. These data were used for further analysis of variance and efficiency.

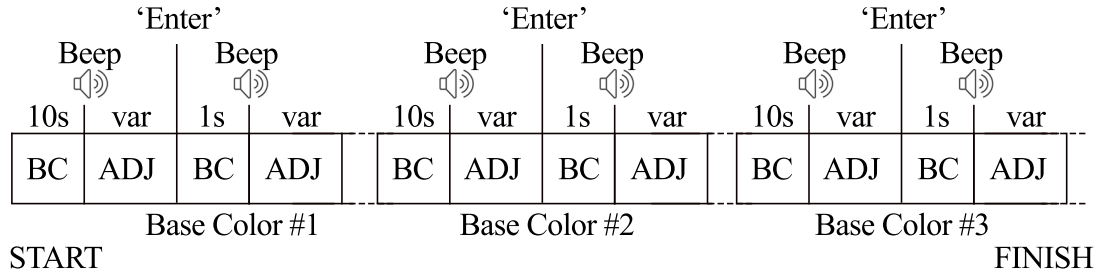


Figure 5.5: Time line for the psychophysical procedure of Method 4 and Method 5; BC stands for base color and ADJ means the adjustment procedure.

5.3.7. ANALYSES

The data of the five methods were analyzed in several consecutive steps as described below.

THRESHOLD CALCULATION

The first step was to get detection thresholds from the data. For both staircase methods (Method 1 and Method 2), the threshold was computed as the average over the last four reversal points. This resulted in 576 threshold values per method (i.e., 3 base colors \times 2 frequencies \times 3 repetitions \times 32 participants). Since the amplitude updating-rule was formulated on a logarithmic scale, the threshold values were expressed as the logarithm of the chromatic modulation in $\Delta(u', v')$. In Method 3, nine stimulus amplitudes were presented ten times for each condition. In order to estimate the threshold of a given condition, a psychometric function was fitted through the (nine) correct response rates as a function of amplitude (on a logarithmic scale) using a generalized linear model (GLM). In particular, we fitted a logistic function:

$$f = \frac{1}{1 + e^{-(a+bx)}} \times 100\% \quad (5.1)$$

where a corresponds to the amplitude at which the response rate is 50%, while b corresponds to the steepness of the psychometric function. Finally, the amplitude corresponding to a probability of 75% was taken as the estimated threshold value. This procedure resulted in 192 (i.e. 3 base colors \times 2 frequencies \times 32 participants) threshold values. Method 4 and Method 5 comprised six adjustments per condition: three times starting from a high amplitude and three times from a low amplitude. For each adjustment, the modulation amplitude of the chromatic flicker stimulus that was presented at the moment that the participant pressed the 'Enter' button, was taken as the estimated threshold value. This resulted in 756 (i.e. 3 base colors \times 2 frequencies \times 2 starting amplitudes \times 3 repetitions \times 21 participants) threshold values.

OUTLIER REMOVAL

Before analyzing the results statistically, we removed outliers at two different levels. At the first level, all the data of a participant for a given method were removed from further analysis when the participant was considered insufficiently sensitive to chromatic flicker or misunderstood the task. This was done only for Method 1, 2 and 3 by taking all responses per participant to the obviously flickering stimulus with a modulation amplitude of 0.05 $\Delta(u', v')$. When a participant indicated less than 80% of the time to see flicker in that stimulus, all the data of that participant for that method were removed. At the second level, the data of a participant was removed only for a specific condition (e.g. base color and/or frequency).

Outliers for a condition were not detected with the most common “mean plus or minus three standard deviations” rule, but were based on the Median Absolute Deviation (MAD), as that method is more accurate to detect outliers [78]. We calculated the corrected Z -score of the threshold averaged across repetitions for each condition, and considered the threshold as an outlier if the absolute value of the corrected Z -score was larger than 2.5, as suggested by Leys et al. [78].

STATISTICAL ANALYSES

After removing the outliers, the following analyses were performed, using the Statistics Toolbox of MATLAB (Release R2015b) and IBM SPSS Statistics for Windows (Version 23).

Reliability of repeated measurements

As mentioned above, participants performed multiple repetitions of the same condition (i.e. base color, frequency and starting amplitude). This resulted in three thresholds per condition for Method 1 and Method 2 (based on the three staircases), one threshold per condition for Method 3 (based on the ten repetitions of the nine amplitudes) and three thresholds per condition for Method 4 and Method 5 (based on the three adjustments). Since each staircase consisted of multiple trials and all staircases were intermingled, there was no well-defined chronological order in the three measures, whereas this was the case for the methods of adjustment. The reliability of the three repeated measurements was evaluated with the Intraclass Correlation Coefficient (ICC). Following the guideline proposed by Koo and Li [63], a single-measurement, absolute-agreement, two-way mixed-effects model was used to calculate the ICCs.

Effect of number of repetitions on the variance of the Detection Threshold

We tested if increasing the number of repetitions would reduce the variance of the detection threshold. Therefore, we calculated the mean threshold per condition in three different ways: (1) based on one measurement, (2) based on two measurements, (3) based on three measurements. For Method 1 and Method 2, the measurements were randomly selected from the three repetitions, because they did not have a well-defined chronological order. However, this random selection is only allowed if the reliability of the three measures is high. For Method 4 and Method 5, we used the first repetition, the first two repetitions and the three repetitions, respectively. Next, a Levene test was performed to test if the variance changed for the three estimates of the detection threshold depending on the number of repetitions.

Effect of Base Color, Frequency and Starting Amplitude on the Detection Threshold

We performed a linear mixed model (LMM) analysis on the relationship between the Detection Threshold as dependent variable and the *Base Color*, *Frequency* and *Starting Amplitude* (only in Method 4 and Method 5) as the fixed independent variables. We also included the interaction terms of the fixed variables and a random intercept for the subjects. For the *Detection Threshold* we always used the average over the various repetitions, since anyway that is the best estimate of the threshold available. The resulting p -values were obtained from a Type III sum of squares. Post-hoc analyses with Bonferroni correction were performed for the significant effects wherever possible.

Normality test and clustering analysis

Where relevant, we checked the normality of the data using the Shapiro–Wilk test. A

possible reason for data not being normally distributed may be different response behavior between participants. We performed hierarchical clustering analyses to check whether participants could be grouped in clusters depending on their threshold data. Here we excluded the participants that were defined as outliers (at the first level), but since the hierarchical clustering algorithm could not handle missing values, the outliers based on a single condition of a participant were included. Two hierarchical clustering methods with a different distance metric were used: (1) Euclidean distance, and (2) linear correlation distance. Where clusters were found, we repeated the LMM analyses as described above for each cluster.

5.4. RESULTS

In this section, we first report the results for each method. In the last sub-section, we statistically analyze the differences between the five methods.

5.4.1. RESULTS OF METHOD 1 – CLASSICAL STAIRCASE WITH YES-NO TASK

When checking the data on outliers according to the procedures described in Section 5.3.7, we removed all (six) threshold values of one participant based on the average response to the maximum amplitude. In addition, we removed seven threshold values of four participants based on the corrected Z -scores. So, in total, thirteen threshold values (i.e. 6.8%) were excluded as outliers. For the remaining thresholds used in further analyses, the corrected Z -scores ranged between -2.11 and 2.47.

Figure 5.6 shows the detection thresholds expressed in $\log_{10} \Delta(u', v')$ for the three base colors and the two frequencies. It illustrates that the threshold slightly decreases for increasing color temperature and increases with increasing frequency, especially for the highest color temperature.

We performed an LMM analysis (as described in Section 5.3.7) to test if *Base Color* and *Frequency* had a significant effect on the detection threshold. The analysis revealed a significant effect of *Base Color* ($F(2, 147.752) = 46.137, p < 0.001$) and *Frequency* ($F(1, 147.901) = 23.233, p < 0.001$), but the interaction between *Base Color* and *Frequency* was not significant ($F(2, 147.752) = 0.338, p = 0.713$). Post-hoc analyses with Bonferroni correction showed that the average threshold at 2700K was significantly higher than at 4000K (MD = 0.074, $p = 0.001$) and at 6500K (MD = 0.187, $p < 0.001$), and the threshold at 4000K was significantly higher than at 6500K (MD = 0.113, $p < 0.001$). Thresholds at 2 Hz were significantly lower than at 4 Hz (MD = 0.078, $p < 0.001$).

The Shapiro-Wilk test showed that the data of each condition was normally distributed. Nonetheless, when combining all data, we found three clusters based on the hierarchical clustering analysis with Euclidean distance; these clusters consisted of 15 (Cluster 1), 12 (Cluster 2) and 4 (Cluster 3) participants, respectively. All participants with one or more outliers were part of Cluster 3. No clear clusters were found based on the cluster analysis with linear correlation distances. In order to check if the results were similar for different groups of participants, we performed an LMM analysis for the participants of Cluster 1, Cluster 2 and the combination of Cluster 1 and Cluster 2. Cluster 3 was not taken into account because it contained too few participants. The analyses all revealed a significant effect of *Base Color* and *Frequency* at a 95% confidence level, which corresponds to the results when the data of all participants were used.

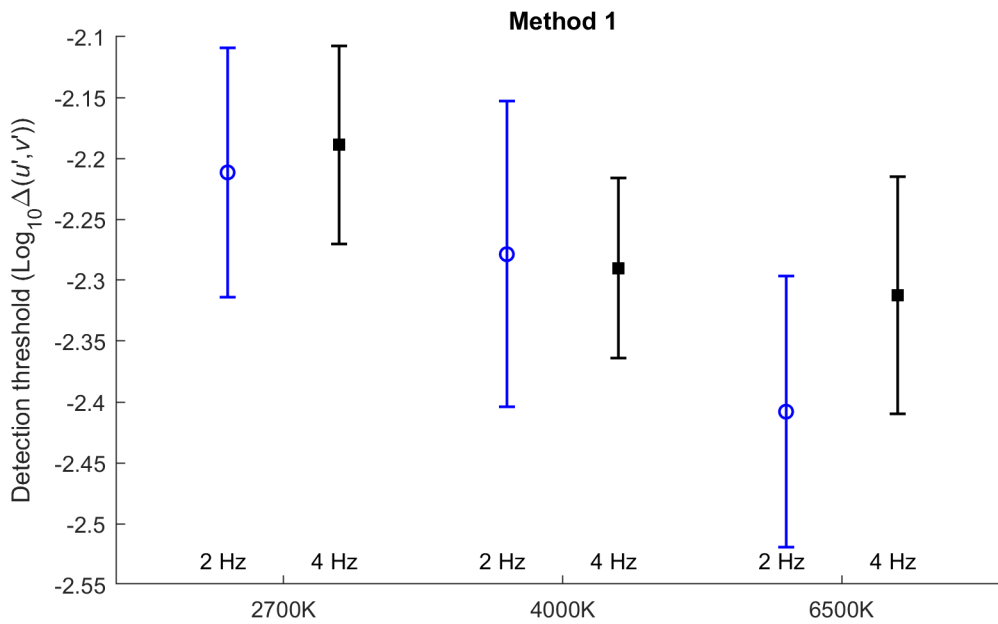


Figure 5.6: Plot of the detection thresholds expressed in $\log_{10} \Delta(u', v')$ of the data (without outliers) obtained with Method 1 for three base colors and two frequencies. The symbols represent the mean threshold values, and the error bars indicate the 95% confidence intervals.

5.4.2. RESULTS OF METHOD 2 – WEIGHTED STAIRCASE WITH 2AFC TASK

In total, thirteen threshold values (i.e. 6.8%) were excluded as outliers: all (six) thresholds of one participant and seven threshold values from five other participants. For the remaining thresholds used in further analyses, the corrected Z -scores ranged between -1.92 and 2.36.

Figure 5.7 shows the detection threshold of the six conditions measured with Method 2. The statistical analysis indicated that *Base Color* had a significant effect on the *Detection Threshold* ($F(2, 148.701) = 19.856, p < 0.001$), while the effect of *Frequency* ($F(1, 148.791) = 0.098, p = 0.754$) and the interaction effect between *Base Color* and *Frequency* ($F(2, 148.867) = 1.504, p = 0.226$) were not significant. Post-hoc analyses showed that the average threshold at 6500K was significantly lower than at 2700K (MD = 0.216, $p < 0.001$) and at 4000K (MD = 0.172, $p < 0.001$). However, the base colors of 2700K and 4000K were not significantly different (MD = 0.044, $p = 0.704$).

The Shapiro-Wilk test showed that the data of all conditions were normally distributed. Two clusters could be defined based on the hierarchical cluster analysis with Euclidean distances, which consisted of 14 (Cluster 1) and 17 (Cluster 2) participants, respectively. No clear clusters were found based on the analysis with linear correlation distances. Additional LMM analyses revealed a significant effect of *Base Color* for each of the two clusters, which is in agreement with the results including all participants. In addition, a significant interaction effect between *Base Color* and *Frequency* was found for Cluster 1 ($F(2, 70) = 3.752, p = 0.028$), showing that the threshold increased with frequency for 2700K and 4000 K, but decreased for 6500K.

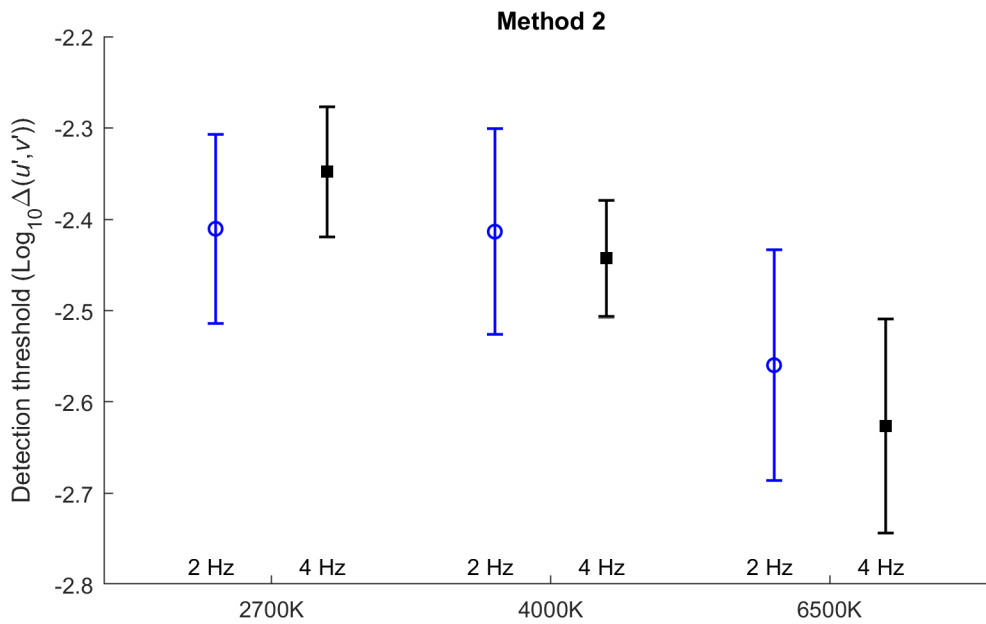


Figure 5.7: Plot of the detection thresholds expressed in $\log_{10} \Delta(u', v')$ of the data (without outliers) obtained with Method 2 for three base colors and two frequencies. The symbols represent the mean threshold values, and the error bars indicate the 95% confidence intervals.

5.4.3. RESULTS OF METHOD 3 – METHOD OF CONSTANT STIMULI WITH 2AFC TASK

Four participants had a correct response rate lower than 80% to the stimulus with the maximum amplitude. Therefore, their data was removed from further analysis. In addition, the data of one other participant had to be removed due to a technical problem. For the remaining twenty-seven participants, we checked if their response to the catch stimulus (with an amplitude of 0) was different from chance. A two-tailed t-test showed that these percentages were not significantly different from 50% ($t(26) = -0.044$, $p = 0.965$), which means that participants had no clear response bias.

For each participant and each condition, the correct response rates of the nine amplitudes were fitted with a GLM model in MATLAB (using **Equation 5.1**). The goodness of the fit can be represented by the deviance of the fit. A deviance of 0 indicates a perfect fit and a high deviance means that the model does not accurately describe the data. In four cases, the deviance was zero. However, these fits were not accurate at all, since there was only one data point that differed from 0% and 100%. Therefore, these four datasets were removed from further analysis. In other cases, the deviance was relatively high and the fit was visually quite bad. In order to decide which values of the deviance were acceptable, we used the criteria that the deviance should follow a normal distribution. As a result, 35 datasets were removed. The Shapiro-Wilk test showed that the remaining 123 deviance values followed a normal distribution ($M = 13.14$, $SD = 6.46$) ($p = 0.054$). For these 123 datasets, the detection threshold was calculated from the fitted psychometric function. Next, fifteen thresholds were removed as outliers based on corrected Z-scores. So, in the end, 108 thresholds with corrected Z-scores ranging from -2.45 to 2.19 were used for analyzing the effect of *Base Color* and *Frequency*.

In Figure 5.8 the thresholds are plotted for the six conditions measured with Method 3. The LMM analysis showed that there was only a significant effect of *Base Color* ($F(2, 85.634) = 14.184, p \leq 0.001$) on the *Detection Threshold*. The effect of *Frequency* ($F(1, 85.505) = 2.694, p = 0.104$) and the interaction effect between *Base Color* and *Frequency* ($F(2, 87.441) = 2.744, p = 0.070$) were not significant. The Post-hoc analyses revealed that the average threshold at 2700K was significantly higher than at 4000K (MD = 0.153, $p = 0.012$) and at 6500K (MD = 0.276, $p \leq 0.001$). The threshold at 4000K was also significantly higher than at 6500K (MD = 0.124, $p = 0.044$).

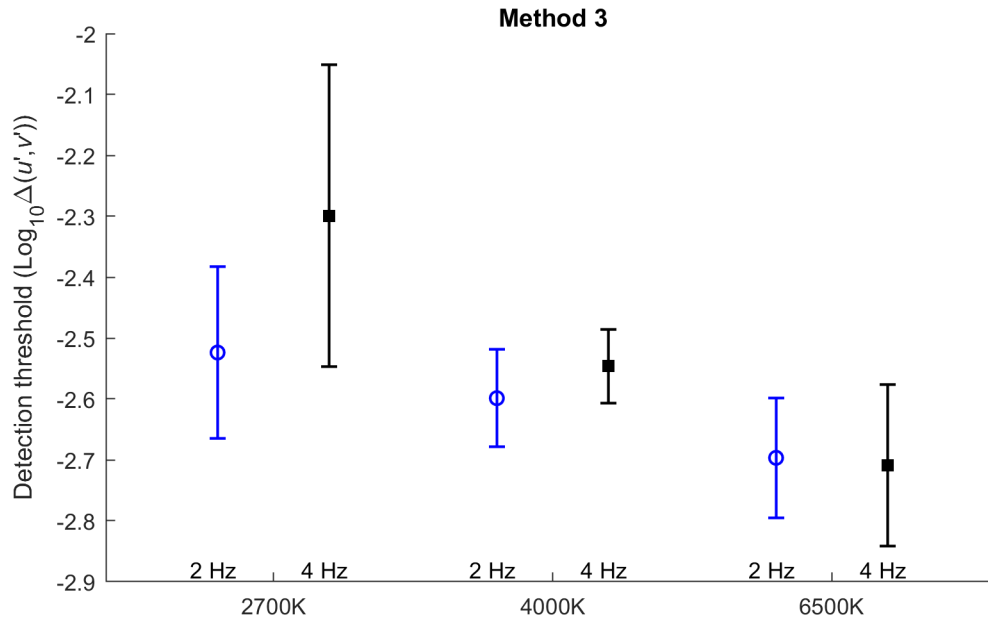


Figure 5.8: Plot of the detection thresholds expressed in $\log_{10} \Delta(u', v')$ of the data (without outliers) obtained with Method 3 for three base colors and two frequencies. The symbols represent the mean threshold values, and the error bars indicate the 95% confidence intervals.

The Shapiro-Wilk test showed that the 108 thresholds were normally distributed under all conditions. Since a relatively large part of thresholds was insufficiently reliable for various reasons (as explained above), we decided not to do a clustering analysis in this case.

5.4.4. RESULTS OF METHOD 4 – METHOD OF ADJUSTMENT WITHOUT REFERENCE

Fifteen threshold values (i.e. 6.0%) from 5 participants were removed as outliers based on the corrected Z -scores, 7 of which came from one particular participant. For the remaining thresholds used in further analyses, the corrected Z -scores ranged from -2.47 to 2.40.

Figure 5.9 presents the detection threshold averaged across participants for the three base colors, the two frequencies and the two starting amplitudes obtained with Method 4. An LMM analysis revealed a significant effect of *Base Color* ($F(2, 215.781) = 5.851, p = 0.003$), *Frequency* ($F(1, 215.830) = 13.085, p \leq 0.001$) and *Starting Amplitude* ($F(1, 215.868) = 4.777, p = 0.030$) on the *Detection Threshold*. None of the interaction effects was significant. Post-hoc analyses showed that the average detection threshold at 6500K was significantly lower than at 2700K (MD = 0.111, $p = 0.004$) and at 4000 K (MD = 0.080, $p = 0.049$). The detection

thresholds at 2700K and 4000K were not significantly different ($p = 1.000$). Thresholds at 2 Hz were significantly lower compared to 4 Hz (MD = 0.099, $p \leq 0.001$). Finally, a low starting amplitude resulted in a lower detection threshold compared to a high starting amplitude (MD = 0.060, $p = 0.030$).

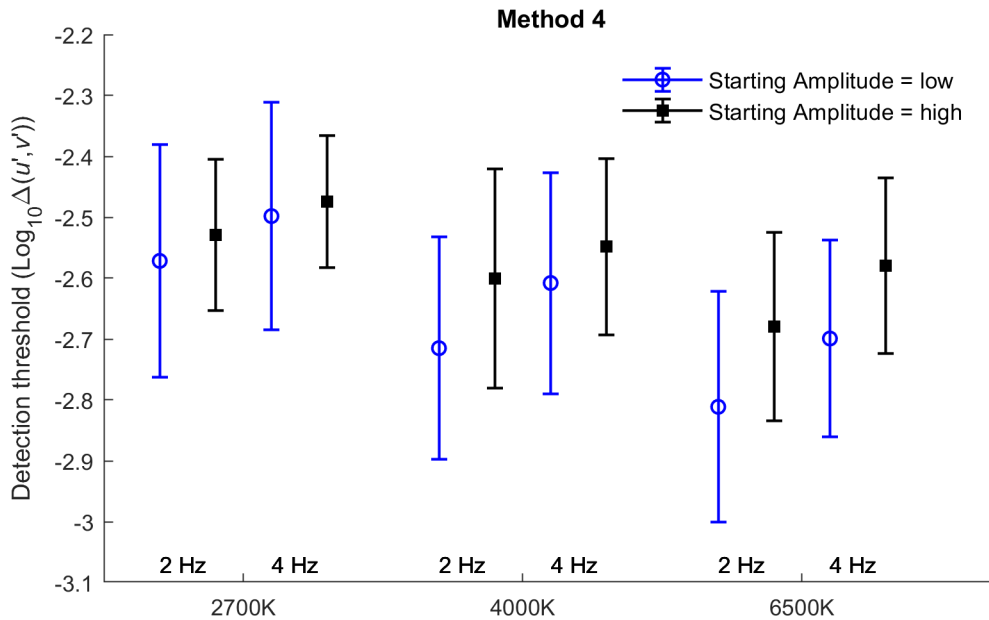


Figure 5.9: Plot of the detection thresholds expressed in $\log_{10} \Delta(u', v')$ of the data (without outliers) obtained with Method 4 for three base colors, two frequencies and two starting amplitudes. The symbols represent the mean threshold values, and the error bars indicate the 95% confidence intervals.

The Shapiro-Wilk test showed that the data under all conditions were normally distributed. Three clusters could be defined based on a hierarchical clustering analysis with Euclidean distances, which consisted of 9 (Cluster 1), 7 (Cluster 2) and 5 (Cluster 3) participants, respectively. By inspecting the raw data, we found that the five participants of Cluster 3 were quite inconsistent and had very low threshold values when the starting amplitude was low. Moreover, four of the five participants in Cluster 3 had one or more outliers. No clear clusters were found based on the analysis with linear correlation distances. Additional analyses were performed on the data of Cluster 1 and Cluster 2. Cluster 3 was not taken into account because it contained only 5 participants. For Cluster 1, only Frequency significantly affected the *Detection Threshold* ($F(1, 99) = 4.309, p = 0.041$). On the other hand, for the participants of Cluster 2, both *Base Color* ($F(2, 75.007) = 20.899, p \leq 0.001$) and *Frequency* ($F(1, 74.839) = 7.806, p = 0.007$) had a significant effect. This was also the case when the participants of Cluster 1 and Cluster 2 were combined (*Base Color*: $F(2, 174.019) = 7.108, p = 0.001$ and *Frequency*: $F(1, 173.991) = 9.580, p = 0.002$). This means that the effect of *Starting Amplitude* disappeared when the participants with inconsistent results were removed.

5.4.5. RESULTS OF METHOD 5 – METHOD OF ADJUSTMENT WITH REFERENCE

Twenty-four threshold values (i.e. 9.5%) from 5 participants were removed as outliers based on the corrected Z-scores. One participant was found to be an outlier in all the 12 conditions.

By checking the raw data of this participant, we found that the modulation amplitudes when the participant pressed ‘Enter’ were all very close to the system’s lower boundary. This participant had the most outliers in Method 4 as well. For the remaining thresholds used for further analysis, the corrected Z -scores ranged from -2.38 to 2.14.

Figure 5.10 presents the detection threshold for the three base colors, the two frequencies and the two starting amplitudes measured with Method 5. The LMM analysis revealed a significant effect of *Base Color* ($F(2, 207.899) = 8.535, p <= 0.001$), *Frequency* ($F(1, 207.998) = 24.994, p <= 0.001$) and *Starting Amplitude* ($F(1, 207.891) = 4.478, p = 0.036$) on the *Detection Threshold*. No significant interaction effect was found. The Post-hoc analyses showed that the average threshold at 6500K was significantly lower than at 2700K (MD = 0.115, $p = 0.001$) and at 4000K (MD = 0.099, $p = 0.004$). Thresholds at 2700K and 4000K were not significantly different ($p = 1.000$). Thresholds were significantly lower at 2 Hz compared to 4 Hz (MD = 0.124, $p <= 0.001$). Finally, a low starting amplitude resulted in a lower detection threshold compared to a high starting amplitude (MD = 0.052, $p = 0.036$).

5

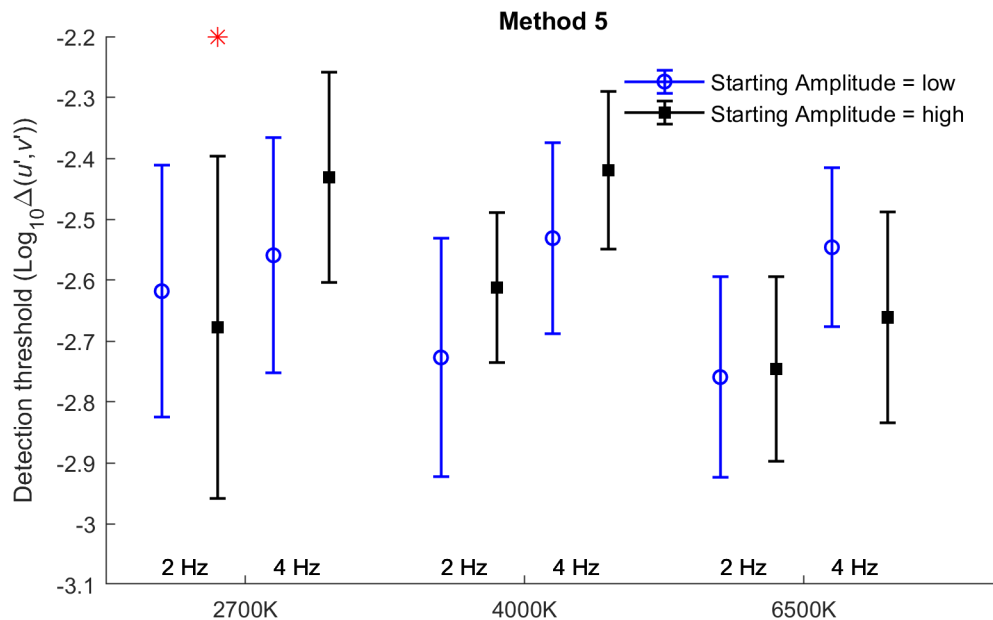


Figure 5.10: Plot of the detection thresholds expressed in $\log_{10} \Delta(u', v')$ of the data (without outliers) obtained with Method 5 for three base colors, two frequencies and two starting amplitudes. The symbols represent the mean threshold values, and the error bars indicate the 95% confidence intervals. The asterisk indicates that the data of this condition was not normally distributed.

The Shapiro-Wilk test showed that one condition was not normally distributed (as indicated by the asterisk in Figure 5.10). Three clusters could be defined based on the analysis with Euclidean distances, and the resulting clusters consisted of 9 (Cluster 1), 8 (Cluster 2) and 3 (Cluster 3) participants, respectively. All participants of Cluster 3 had one or more outlying values in their measured data. They also had very low threshold values when the starting amplitude was low. No clear clusters were found based on the analysis with linear correlation distances. For Cluster 1, *Base Color* ($F(2, 99) = 29.377, p <= 0.001$) and *Frequency* ($F(1, 99) = 12.989, p <= 0.001$) significantly affected the *Detection Threshold*. For Cluster 2, the effect of *Frequency* ($F(1, 87.010) = 14.335, p <= 0.001$) and *Starting Amplitude*

	ICC	95% CI	Result of Levene test
Method 1	0.930	0.902-0.952	$F(2, 261) = 0.112, p = 0.894$
Method 2	0.293	0.161-0.431	$p > 0.05$ (85% of the case)
Method 4	0.686	0.618-0.746	$F(2, 522) = 0.557, p = 0.573$
Method 5	0.714	0.650-0.771	$F(2, 534) = 0.512, p = 0.600$

Table 5.3: Results of the reliability test over repetitions per method in terms of ICC and their 95% confidence interval (CI). The last column gives the results of a Levene test to check whether the variance depended on the number of repetitions

($F(1, 87.010) = 4.339, p = 0.040$) were significant. For the combination of Cluster 1 and Cluster 2, *Base Color* ($F(2, 185.996) = 17.176, p \leq 0.001$), *Frequency* ($F(1, 185.996) = 23.873, p \leq 0.001$) and *Starting Amplitude* ($F(1, 185.996) = 5.058, p = 0.026$) significantly influenced the *Detection Threshold*. This indicates that not for all participants the measured threshold depended on the starting amplitude.

5.4.6. COMPARISON OF METHODS

In the previous sections (5.4.1 to 5.4.5), the experimental results are shown per method. However, as mentioned before, we were mainly interested in comparing the methods in terms of reliability, precision, accuracy and efficiency. Therefore, additional analyses were performed on the combined data of all methods. Since we found large between-subject differences in our previous analyses, we decided to only use the data of those participants who performed all five methods and were not marked as an outlier in any of these methods. As a result, the data of fifteen participants were used for the following analyses.

RELIABILITY

A psychophysical method should be reliable, meaning that the same results are obtained by repeating the experiment. The reliability of the repeated measurements was evaluated with the Intraclass Correlation Coefficient (ICC). Note that Method 3 aimed at measuring the entire psychometric curve and resulted in only one threshold. Hence, this method was not included in the reliability analysis. The ICCs of the other four methods are summarized in Table 5.3. The table illustrates that the reliability was excellent for Method 1, and moderate for Method 4 and Method 5. When calculating the average standard deviation between the three measurements, we found a reasonably small value of 0.06 for Method 1, and 0.13 for Method 4 and Method 5. For Method 2, however, the ICC revealed poor reliability. In this case, the average standard deviation between the measurements was also considerably larger (i.e., 0.25).

EFFECT OF NUMBER OF REPETITIONS ON THE VARIANCE OF THE DETECTION THRESHOLD

A psychometric method is considered as efficient when the variance does not decrease significantly when the number repetitions, and hence also the number of trials, increases. Figure 5.11 shows per method the mean and 95% confidence interval of the detection threshold (averaged across all conditions and all participants) for different number of repetitions. Again note that for Method 3 we only have one “repetition”. For Method 1, we calculated the detection threshold by averaging over (1) one randomly selected repetition, (2) two

randomly selected repetitions, and (3) all three repetitions, which was allowed based on the high ICC. Since the ICC for Method 2 revealed poor reliability, this procedure would, however, not yield an accurate representation of the mean and variance. Therefore, we repeated the procedure of randomly selecting one, two or three repetitions multiples times ($N = 1000$) and we plotted the mean confidence interval in Figure 5.11. For Method 4 and Method 5 we averaged over (1) the first repetition, (2) the first two repetitions, and (3) all three repetitions, since these were measured in chronological order. A Levene test revealed that the variance did not change significantly with the number of repetitions for Method 1, Method 4 and Method 5, which is in line with the observation that the confidence intervals in Figure 5.11 are similar for different numbers of repetitions. For Method 2, the effect of repetition was significant only in 15% of the cases, and as can be seen in Figure 5.11, the mean variance appears to decrease with the number of repetitions. The results of all Levene tests are summarized in Table 5.3.

5

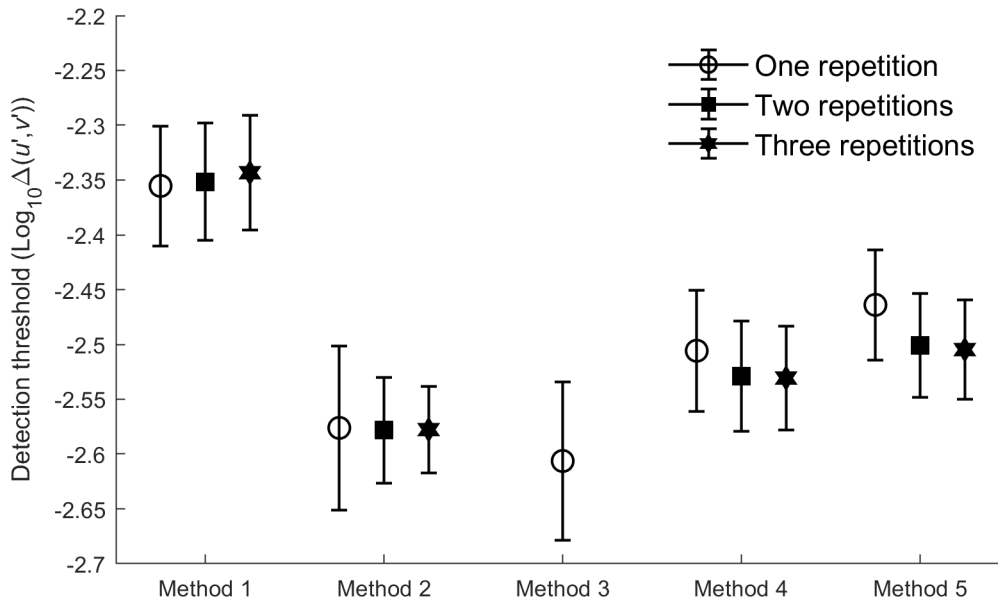


Figure 5.11: The detection thresholds expressed in $\log_{10} \Delta(u', v')$ averaged across all conditions and participants using either one, two or three repetitions. The symbols represent the mean threshold values, and the error bars indicate the 95% confidence intervals. Note that for Method 2, the mean threshold and confidence interval was based on the mean of the 1000 random selections.

EFFECT OF METHOD, BASE COLOR AND FREQUENCY ON THE DETECTION THRESHOLD

Figure 5.12 shows the detection threshold averaged over all conditions and all (fifteen) participants for each of the five methods. An LMM analysis was performed with *Detection Threshold* as dependent variable and *Method*, *Base Color* and *Frequency* as the fixed independent variables, including the interaction terms between these independent variables and a random intercept for the participants. The *Detection Threshold* was averaged over the three repetitions (in all methods except Method 3) and also over the two starting amplitudes in Method 4 and Method 5. The analysis revealed a significant effect of *Base Color* ($F(2, 397.286) = 27.535, p \leq 0.001$), *Frequency* ($F(1, 397.326) = 5.464, p = 0.020$) and *Method* ($F(4, 397.530) = 15.168, p \leq 0.001$). No significant interaction effect was found. The Post-hoc

analyses showed that the average threshold at 6500K was significantly lower than at 2700K (MD = 0.215, $p \leq 0.001$) and at 4000K (MD = 0.119, $p \leq 0.001$); thresholds at 4000K were significantly lower than at 2700K (MD = 0.096, $p = 0.003$). These results are as expected, since this tendency was also found for each of the methods separately. Thresholds significantly increased from 2 Hz to 4 Hz (MD = 0.055, $p = 0.020$). Although this was not significant for each method separately, all methods showed the same trend. Finally, all methods resulted in the same value of the mean detection threshold, except for Method 1. The Post-hoc analysis showed that the detection threshold for Method 1 was significantly higher than for Method 2 (MD = 0.235, $p \leq 0.001$), Method 3 (MD = 0.257, $p \leq 0.001$), Method 4 (MD = 0.186, $p \leq 0.001$) and Method 5 (MD = 0.162, $p \leq 0.001$), respectively. Thus, although all methods revealed the same trend for the base color and frequency, the actual values of the detection threshold were not equal for all methods.

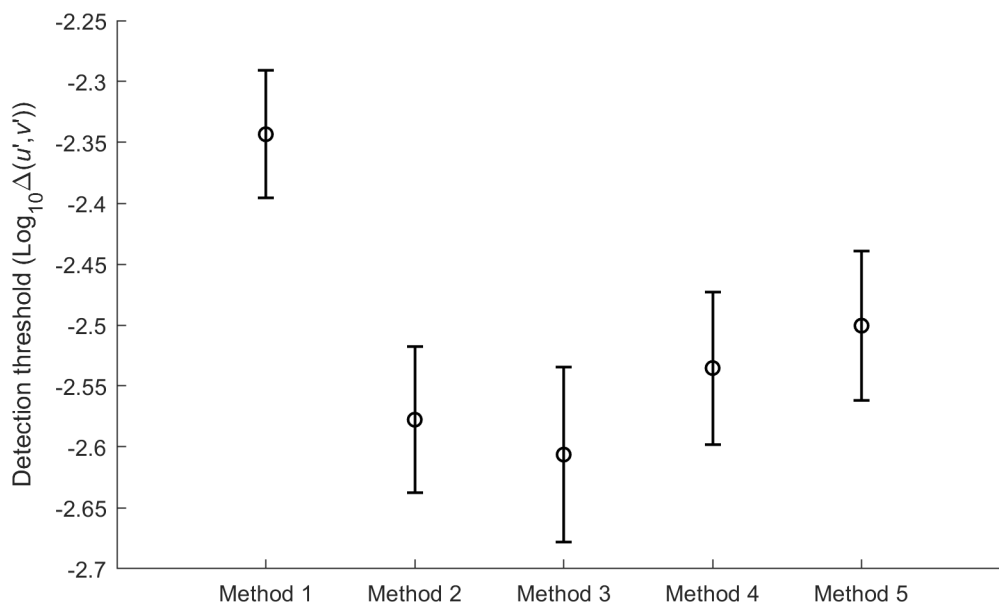


Figure 5.12: The detection threshold expressed in $\log_{10} \Delta(u', v')$ averaged over all conditions and participants for each of the five methods. The symbols represent the mean threshold values, and the error bars indicate the 95% confidence intervals.

ACCURACY OF THE METHODS

As mentioned in Section 5.1, accuracy refers to the deviation of the measured threshold from the veridical or “true” threshold. The veridical threshold can of course never be determined in an empirical experiment. However, since Method 3 not only measured a threshold, but estimated the entire psychometric function, we used the data of Method 3 as the ground truth. In order to estimate the bias, we translated each value of the measured detection threshold into the percentage of the underlying psychometric function for that condition. Since the psychometric functions of Method 3 were based on a 2AFC task (with responses correct varying between 50% and 100% and a threshold at 75%), we first transformed these into psychometric functions for a yes-no task (with responses “yes” varying between 0% and 100% and a threshold at 50%). Then we calculated the percentage “yes” responses on the

corresponding hypothetical veridical psychometric function for the detection threshold of each participant and each condition.

An LMM analysis on the percentage “yes” responses revealed a significant effect of *Method* ($F(3, 229.543) = 7.899, p \leq 0.001$). Additional t-tests revealed that the percentage “yes” responses significantly differed from 50% for Method 1 ($t(60) = 6.647, p \leq 0.001$), Method 4 ($t(60) = 2.600, p = 0.012$) and Method 5 ($t(62) = 2.112, p = 0.039$). Participants indicated that the chromatic flicker was just visible at a probability of 71%, 59% and 57% for Method 1, Method 4 and Method 5, respectively. Note that Figure 5.12 shows the same trend for the mean detection threshold, even though Method 4 and Method 5 were not significantly different from Method 3. However, by transferring the threshold into a percentage “yes” responses on the psychometric function measured in Method 3, the values can be directly compared to a fixed value of 50% and this makes the differences more significant.

5

EFFICIENCY OF THE METHODS

As pointed out in Section 5.1, efficiency is related to the number of trials required to achieve a certain precision. So, if the precision between two methods is the same, one can take the number of trials as a measure of efficiency. As can be seen in Figure 5.12, the variance of all methods was very similar, which was also confirmed by a Levene test ($F(4, 407) = 1.258, p = 0.286$). Since the variance of all methods was the same, we can express the efficiency in terms of the number of trials required to determine a threshold (Figure 5.13(a)), and the time needed to determine a threshold (Figure 5.13(b)). Notice that the threshold was measured three times for the staircase methods and six times for the methods of adjustment (including a high and low starting amplitude). However, since the variance did not change substantially when increasing the number of repetitions, we can base the efficiency estimation on a single measurement of the threshold value. Both graphs of Figure 5.13 show the same trend. As expected, Method 3 had the highest number of trials per threshold (which was chosen to be 90), and also took longest (approximately 296 seconds). At the same time, it also provided more information than only the threshold. The two methods of adjustment were most efficient. Participants used on average 13 trials and took 23 seconds to find the threshold, both for Method 4 and for Method 5. The two staircase methods performed in between, with Method 1 being faster and requiring less trials than Method 2.

5.4.7. QUESTIONNAIRE ANALYSIS

Participants were asked to evaluate their experience at the end of a session. In general, they found the experiment quite difficult for all methods. Participants also thought the experiment was slightly too long, since most of them experienced eye strain at the end of the session. Most participants indicated that the method of adjustment was easier when starting from a high amplitude compared to a low amplitude. In addition, most participants mentioned that the reference without flicker was useful for the method of adjustment, as they felt more sure when pressing the ‘Enter’ button. However, some participants reported that the reference appeared to flicker, which confused them. Since participants had to make eye movements to be able to compare the test and reference stimuli, the apparent flicker of the reference stimulus could be caused by flicker adaptation.

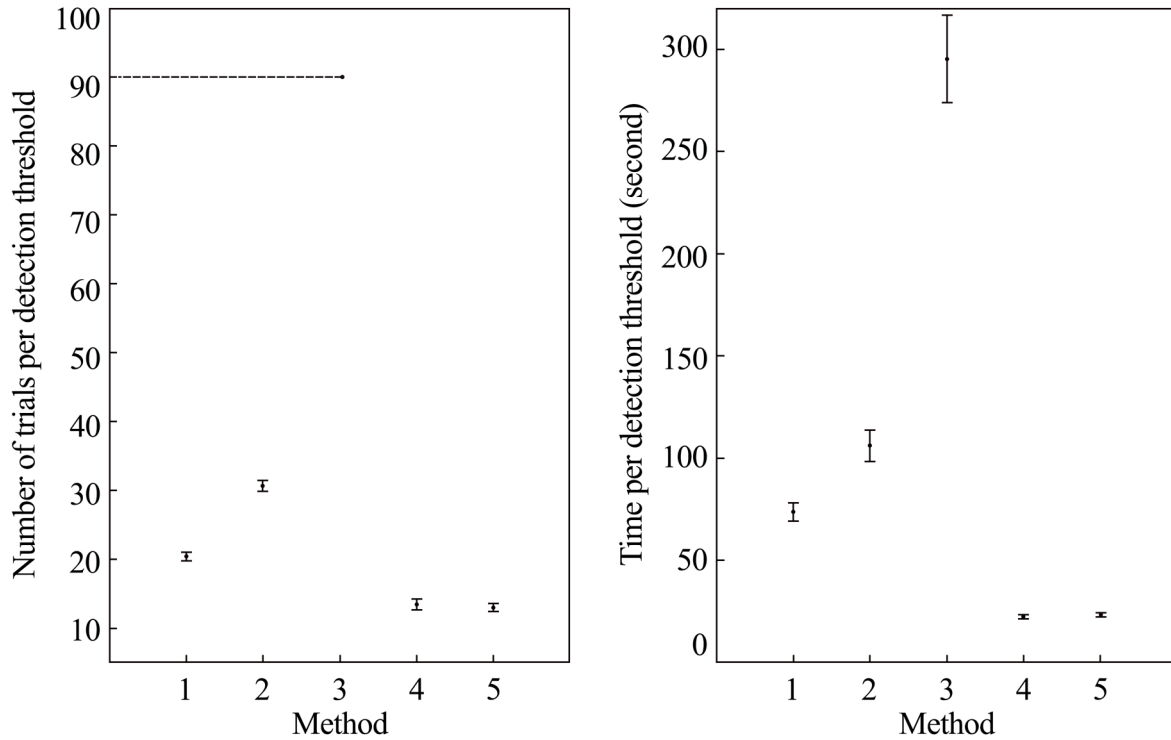


Figure 5.13: (a) The number of trials per threshold for the five methods (Method 3 - the method of constant stimuli had a fixed number of 90 trials); (b) The time per threshold for the five methods (in second)

5.5. DISCUSSION

This study compared the accuracy, precision and efficiency of five commonly used psychophysical methods for measuring the detection threshold of chromatic flicker. The methods employed were the classical 1-up-1-down staircase method with a yes-no task (Method 1); the weighted 3-up-1-down staircase method with a 2AFC task (Method 2); the method of constant stimuli with a 2AFC task (Method 3); the method of adjustment without reference (Method 4), and, the method of adjustment with reference (Method 5). The chromatic flicker stimuli were temporally modulated with a square wave, i.e. two colors were shown alternately, where the mean chromaticity of the two colors was either 2700K, 4000K or 6500K, and the temporal frequency could be 2 Hz or 4 Hz. Below we summarize the main conclusions, formulate recommendations for future experiments on chromatic flicker detection and reflect on possible limitations of the study.

5.5.1. EFFECT OF BASE COLOR AND FREQUENCY

When the data of all experiments were combined, we found that the detection threshold expressed in $\log_{10} \Delta(u', v')$ decreased with increasing color temperature of the base color of the chromatic flicker stimulus. To our knowledge this has not been studied before. It shows that the CIE 1976 UCS color space is not a suitable model to represent the sensitivity of human observers to temporal color differences. Other studies have shown that also in the CIE 1976 $L^*a^*b^*$ color space, temporal color differences are not perceptually uniform [121, 97]. We also found that the detection threshold increased with frequency from 2 Hz to 4 Hz. This suggests that chromatic flicker perception exhibits a band-pass behavior. Literature has

found inconclusive results on the sensitivity to chromatic flicker at relatively low frequencies (e.g., [36, 45, 136]). Some researchers claim that the increase with frequency is actually caused by apparent luminance flicker. Hence, much more data is needed on the influence of frequency, base color and other factors on the perception of chromatic flicker. This can in the end be used to develop an accurate and useful temporal color space, which is essential to design pleasant dynamic light. However, before collecting a lot of data, the most efficient methodology has to be selected.

5.5.2. COMPARISON OF METHODS

The comparison of the five psychometric methods in terms of accuracy, precision and efficiency yields several recommendations for future research.

ACCURACY

5

For Method 1, Method 4 and Method 5 the measured threshold deviated from a detection probability of 50%, assuming that the psychometric function of Method 3 represented the ground truth. The estimated decision criterion was larger than 50%, which implies that the observers wanted to be more “confident” to indicate that flicker was visible. However, it is not known how accurate Method 3 estimated the 50% detection threshold. Literature indicates that a psychometric function can be shifted towards the center of the chosen intensity levels [76], for instance, because participants tend to equalize the frequencies of the (two) response categories [119]. Nevertheless, the detection threshold of Method 3 corresponded to the threshold measured in Method 2. This very promising, because the task of the participant was the same in these two methods but the stimuli that were presented could be quite different.

We hypothesized before that the mean threshold of Method 1 would be equal to that of Method 4 and Method 5, because in all these methods participants were asked to indicate if they saw the difference or not. However, the yes-no staircase method revealed a higher threshold. In Method 4 and Method 5, participants knew what the next stimulus would be since they changed it themselves, whereas in the staircase method the next stimulus was unknown as it was determined by the system. This might have influenced the strategy that participants used in their judgement.

The thresholds obtained with an adjustment task (Method 4 and Method 5) were slightly higher compared to the thresholds obtained with a 2AFC task (Method 2 and Method 3), as can be seen in Figure 5.12. This was also expected, since the method of adjustment depends more on the decision criterion of the participant than the 2AFC method, where participants are forced to judge which of the two stimuli are flickering. However, the differences were not significant when taking into account the variance between participants. It is, therefore, interesting that in the case of chromatic flicker perception both methods yield approximately the same detection thresholds.

PRECISION

For Method 1, 4 and 5 the precision of the detection threshold did not improve by averaging over repeated measurements. This means that the duration of the experiment can be reduced by measuring the threshold only once. It also indicates that the variance between participants is larger compared to the variance within participants. Obviously, the differences between participants cannot be reduced by measuring the threshold per participant more often. It has been reported before that there are large individual difference in

chromatic contrast sensitivity [38].

The reliability of Method 2 was relatively poor and the variance slightly decreased with more repetitions. The within-subject variability of this method was also larger compared to the other methods. We noticed that some staircases converged to an unrealistic small or large value. This was also evident from the outliers, detected through the corrected Z -scores of the thresholds. It is not clear what caused this large within-subject variability. It may be related to the nature of the weighted 3-up-1-down staircase method or to the 2AFC task. Method 1 also used a staircase method, but a different up/down rule. Method 3 also used a 2AFC task, but here we didn't measure the variability within subjects.

We found that the variance did not significantly differ between the five methods when all repeated measurements were taken into account. This is contradictory to the hypothesis formulated in Section 5.2.4. We expected that the variance would be smallest when the effect of decision criterion was minimized, a reference stimulus was presented and the participants did not have control over the stimuli. All these conditions were fulfilled in the two methods with a 2AFC task. Apparently, the variance caused by these factors was negligible compared to the difference in chromatic flicker sensitivity between participants. This means that for all methods the variance was mainly dominated by the (same) between-subjects variance. This could also explain the finding that the mean value of the detection threshold did not change with number of repetitions for any of the methods.

EFFICIENCY

We found that the method of constant stimuli (Method 3) is the most inefficient method for determining the detection threshold of chromatic flicker. It needs a substantially larger number of trials and more time, which may only be worthwhile to invest if one wants to measure the whole psychometric curve. Apart from the larger time investment in the experimental execution, it also requires considerable effort upfront in estimating appropriate intensity levels of the stimuli. Method 2 is the second most inefficient, followed by Method 1, while both Method 4 and Method 5 are the most efficient methods. This is largely in line with the hypothesis formulated in Section 5.2.4.

Obviously, the efficiency may depend on the design parameters for the experimental implementation of the different methods. Changing the number of reversal points in a staircase, or the step sizes in the adjustment method will affect the number of trials and time needed to measure the threshold. However, since these design parameters may also impact the variance of the measured threshold, it is hard to predict how they will affect efficiency in the end.

SUMMARY

If an accurate estimate of the mean detection threshold of chromatic flicker is required and time is not an issue, one could use the method of constant stimuli with a 2AFC task or the weighted 3-up-1-down staircase method with a 2AFC task. The method of adjustment with or without a reference is also a good candidate, but might yield slightly higher values. When the staircase method with a 2AFC task is being used, it is recommended to use two or more repetitions in order to reduce the variance of the threshold. When the method of constant stimuli is being used, one should carefully select the range of modulation amplitudes. Otherwise, the fitted logistic function might become unreliable for some conditions, as was the case in our experiment.

However, when time is limited or when the detection threshold of chromatic flicker needs to be determined for many different conditions, the method of adjustment without reference is favorable. It is a factor 3 to 4 more efficient than the staircase methods, it yields a threshold that is not significantly different from the threshold measured with the other methods (except for Method 1), and the variance of the threshold is comparable to that of the other methods. The variance is mainly determined by the relatively larger difference between participants, so, a within-subject design is recommended.

The positive conclusion on the method of adjustment differs from what is known in literature. As mentioned, Wier, Jesteadt, and Green [158] and Podlesek and Komidar [106] concluded that the threshold measured with the method of adjustment was considerably different from the threshold measured with a staircase method or a method of constant stimuli, whereas we found very similar thresholds for these methods. In addition, Podlesek and Komidar [106] showed that the variability in threshold was largest for the method of adjustment, whereas it was comparable for all methods in our experiment. On the other hand, our results are in line with the findings of Eisen-Enosh et al. [19], who didn't use the method of adjustment, but showed that the method of constant stimuli and a weighted staircase method obtained high agreement in the measurement of the critical flicker-fusion frequency, while the method of constant stimuli had the lowest efficiency. It is not a priori clear why our conclusions differ from the ones of Wier, Jesteadt, and Green [158] and Podlesek and Komidar [106]. It may be related to the kind of stimulus that is presented (i.e. an auditory tone for Wier, Jesteadt, and Green [158] and a visual target moving in the frontal plane for Podlesek and Komidar [106]), and to the degree of variability among people in the sensitivity to that stimulus.

5.5.3. POSSIBLE LIMITATIONS

FLICKER ADAPTATION

Adaptation is a well-known phenomenon in visual perception, and also flicker adaptation has been described in literature before [122]), at least for stimuli that fluctuate in luminance. Therefore, we can't exclude that also in our experiment with stimuli fluctuating in chromaticity, flicker adaptation occurred. So, our measured thresholds might be affected by attenuations in the observers' flicker sensitivity after prolonged exposure to flickering stimuli. We tried to minimize this effect by adding a reference in some of the methods used, and by using counterbalancing experimental designs where possible. Still, some participants reported to see flicker in the non-flickering reference stimuli, which might be an indication of the influence of flicker adaptation.

SELECTING STIMULUS INTENSITIES FOR THE METHOD OF CONSTANT STIMULI

As mentioned earlier, the method of constant stimuli requires some pre-knowledge to select the right stimulus intensities. Of course, the range of stimulus intensities should be spread around the threshold intensity level, but the problem is that the threshold is actually unknown. In literature, the stimuli intensities usually are chosen on the basis of a model or pilot studies [76]. In our case, we first roughly analyzed the data of both staircase methods (Method 1 and Method 2) to determine the range of modulation amplitudes for the method of constant stimuli (Method 3). Since a first look at the data revealed large individual differences, we estimated the right stimulus range per participant, by averaging over the three base colors and the two frequencies. The latter might have influenced the accuracy of the

selected stimulus intensities, since the detection threshold for all methods depended on the base color, and in most cases also on the frequency. Moreover, we combined the data of Method 1 and Method 2, while after careful analysis we found a significant difference between Method 1 and the other methods. This also made the selection of stimulus intensities less accurate. Indeed, the fit of the psychometric function was not acceptable in many of the cases, so we had to remove almost half of the data. However, since we found – in the end – the same threshold for all methods except for Method 1, we can conclude that the threshold of Method 3 was measured with sufficient accuracy.

5.6. CONCLUSIONS

This study investigated the accuracy, precision and efficiency of five commonly used psychophysical methods for measuring the detection threshold of chromatic flicker. The classical 1-up-1-down staircase method with a yes-no task (Method 1) was very reliable, the variance was comparable to the other methods and did not improve with repeated measurements, the efficiency was moderate, but the detection threshold was higher compared to the other methods. The weighted 3-up-1-down staircase method with a 2AFC task (Method 2) was less reliable and had the largest within-subject variability, the variance slightly decreased when two repeated measurements were used, the efficiency was moderate and the detection threshold was similar to three of the other methods. The method of constant stimuli with a 2AFC task (Method 3) was very inefficient, it took about 10 times longer than the most efficient one and it also had the largest number of data that had to be removed because of inaccuracies. The two methods of adjustment with (Method 4) and without a reference (Method 5) were quite reliable, the variance was comparable to the other methods and did not improve with repeated measurements, the efficiency was extremely high, and, the detection threshold was similar to that of the methods with a 2AFC task. The method of adjustment without a reference yielded slightly more accurate thresholds than the method of adjustment with a reference compared to the assumed ground truth data. Therefore, it is recommended to use the method of adjustment without reference in combination with a within-subject design in future studies to measure the detection threshold of chromatic flicker.

The experiment also enabled us to explore the effect of the base color and frequency of the chromatic flicker stimulus on the detection threshold expressed in $\log_{10} \Delta(u', v')$. The results demonstrated that people are more sensitive to chromatic flicker at 4 Hz compared to 2 Hz and their sensitivity increases with color temperature. The latter implies that CIE 1976 UCS color space is not a useful model for temporal color perception. In order to develop an accurate model, more data is required on the perception of temporally modulated colored light.

6

Modeling Contrast Sensitivity for Chromatic Temporal Modulations

The temporal contrast sensitivity to isoluminant chromatic flicker was measured for three observers using the method of adjustment. The isoluminant stimuli were created for each observer individually, based on a technique similar to heterochromatic flicker photometry (HFP). The chromatic flicker stimuli were sinusoidal modulations, defined in the CIE 1976 UCS (u', v') chromaticity diagram. The chromaticity varied around a base color along a certain modulation direction with a certain amplitude at a certain frequency. Nine base colors, four modulation directions and seven frequencies were used, resulting in thirty-six temporal contrast sensitivity curves per observer. An exponential model was fitted to the resulting contrast sensitivity expressed as $1/\Delta(u', v')$, $1/\Delta LMS$ and $1/\Delta lms$. The model resulted in an average R^2 value higher than 0.93 for the three different measures of contrast sensitivity. The two parameters of the model (i.e. the slope β_1 and intercept β_0) were found to significantly depend on the base color and direction of the chromatic modulation. This means that $\Delta(u', v')$, ΔLMS and Δlms are not suitable measures to predict the sensitivity to temporal chromatic modulations in different locations of the color space.

This chapter is copied with (slight) adaptations from The Twenty-sixth Color and Imaging Conference (CIC) Final Program and Proceedings as: "Modelling Contrast Sensitivity for Chromatic Temporal Modulations", Xiangzhen Kong, Mijael R. Bueno Pérez, Ingrid M.L.C. Vogels, Dragan Sekulovski and Ingrid Heynderickx [60].

6.1. INTRODUCTION

The development of light emitting diodes (LEDs) technology has enabled inexpensive ways to easily create dynamic colored light. However, current knowledge on human perception of dynamic colored light is limited and still insufficient to provide guidelines for comfortable and attractive implementations. Some studies aiming at understanding the perception of dynamic light have concentrated on the perceived smoothness [121], the preference [145], and the perceived subtlety [97] of temporal color transitions. The limited number of studies show that existing spatial color spaces cannot be used to accurately predict these phenomena and that a temporally uniform color space is needed. The sensitivity to temporal modulations in luminance and chromaticity is a useful paradigm to collect data for building such a model.

The sensitivity of the human visual system to temporal modulations is known to depend on the modulation frequency. Above a certain critical fusion frequency (CFF), the modulation cannot be perceived independent of its magnitude. Below the CFF, the relationship between temporal frequency and sensitivity is called the Temporal Contrast Sensitivity Function (TCSF). TCSFs have been extensively studied in the past mainly for two purposes. First, in clinical vision science, for instance, to detect a variety of pathologies affecting the visual system [13, 103]. Second, to understand the underlying mechanisms of human vision [110, 141, 136, 64, 20, 93, 164, 49, 132].

In order to build a temporal color space a large amount of data has to be collected over the entire color space. Therefore, an efficient way of sampling the color space is required. As a first step, it would be beneficial to have a model describing the TCSFs with a small number of parameters, so we can use less temporal frequencies to subsequently study the TCSF in various locations of the color space. Researchers have found that the TCSF for luminance modulations is generally a band-pass function [15]. For chromatic modulations the TCSF is usually a low-pass function, where the sensitivity decreases with increasing frequency [44, 16]. However, some studies have found a small decrease in sensitivity also at low frequencies [81]. The TCSF has also been found to depend on the demographics (e.g., age) of the observers as well as stimulus features (e.g., stimulus size, retinal illuminance, characteristics of the surrounding) [136, 159].

Several models for TCSFs have been proposed. For instance, Dobkins, Lia, and Teller [16] used a double exponential function to fit the TCSFs for luminance and found that a single exponential function was sufficient to describe the TCSF for chromaticity. Other physiologically based models exist based on stimuli that activate only part of the visual system (e.g., the red-green opponency channel) [20]. Due to technical limitations at the time, most experiments on chromatic TCSFs have been carried out for limited color stimuli, usually red-green chromatic flicker. In this study, we investigate if the relationship between chromatic contrast sensitivity and temporal frequency can be described by the same exponential function for a wide range of chromatic flicker stimuli. In particular, we vary the base colors (i.e. the mean color of the chromatic modulation) and the modulation direction in the 1976 UCS (u', v') chromaticity diagram.

6.2. METHOD

In this study, the detection threshold of chromatic flicker was measured for 9 base colors, 4 directions of chromaticity change and 7 temporal frequencies, using a full-factorial within-

subject design with 3 participants. Before the main experiment, a preparation experiment was performed to determine the luminance ratios between two alternating colored stimuli that minimizes the visibility of luminance flicker for a given participant using a method similar to heterochromatic flicker photometry (HFP). The luminance ratios were determined for the 36 color pairs used in the main experiment with a chromaticity difference of $0.05 \Delta(u', v')$ at a flicker frequency of 25 Hz. The individual luminance ratios were used to make the isoluminant chromatic flicker stimuli of the main experiment.

6.2.1. EXPERIMENTAL SETUP

A specially designed LED system, with 36 Cree XP-E LEDs arranged in a square panel (12 red, 8 green and 16 blue LEDs), was used. The system was calibrated with a spectrometer and shown to be reliable and stable over time. The LEDs were driven by means of pulse width modulation (PWM) at a driving frequency of 2 kHz and having 16-bit dimming. The PWM signals were generated by an Arduino Due microcontroller, which was connected to a lab computer. The drivers of the LEDs accepted RGB values, in the device dependent color space of the LEDs. The target stimuli were defined in CIE 1976 UCS (u', v') and transformed via XYZ to the RGB values of the LEDs.

The LED panel was placed in a box (height: 1.5 m, depth: 0.8 m, width: 1.5 m) with a circular opening of 26.4 cm in diameter (see Figure 6.1). Participants were seated at a predefined position of 1.5 m from the front of the box, which resulted in a stimulus covering a visual angle of 10-degrees. Since we are interested in using the data for realistic lighting applications, a 10-degrees field was preferred over 2-degrees. Participants could only see the stimuli from this circular opening. The inner surfaces of the box were smooth and colored natural white. The LEDs were mounted in such a way that the visible light field was quite uniform (i.e., the luminance deviated by a maximum of 3.3% at the stimuli luminance level of 37.5 cd/m^2 , while the chromaticity deviated by a maximum $\Delta(u', v') = 0.0013$). The inner edges of the box and the luminaries were not visible.

To avoid head movements of the participants, a chinrest was used. Additionally, a standard keyboard was provided as an input device for the participants.

6.2.2. STIMULI

The stimuli consisted of light sinusoidally modulated in time. The chromaticity of the light varied around a base color in a predefined direction specified in the CIE 1976 UCS (u', v') chromaticity diagram. Depending on the individual luminance ratios, the luminance of the light varied at the same temporal frequency as the chromaticity to make the stimulus isoluminant for each participant.

Nine base colors (BC_1 to BC_9) were selected (see Figure 6.2). The chromaticities of BC_1 , BC_2 and BC_3 were close to the green, blue and red LED, respectively. BC_4 , BC_5 and BC_6 were the middle points between BC_1 and BC_3 , BC_1 and BC_2 , BC_2 and BC_3 respectively, while BC_7 was located in the center of the gamut. During the experiment we discovered that it was impossible to make the entire 10-degree visual field non-flickering for BC_2 and BC_3 at the selected frequencies and amplitudes. Instead, the colors could be fused either for the center of the field or at the outer edge, but not for the entire field at the same time. The possible explanation is that for BC_2 and BC_3 , the contribution of high-energy blue light was high and the influence of macular pigment concentrated in the fovea might have contributed to non-uniform flicker visibility over the 10-degrees field of view. Therefore, two other base

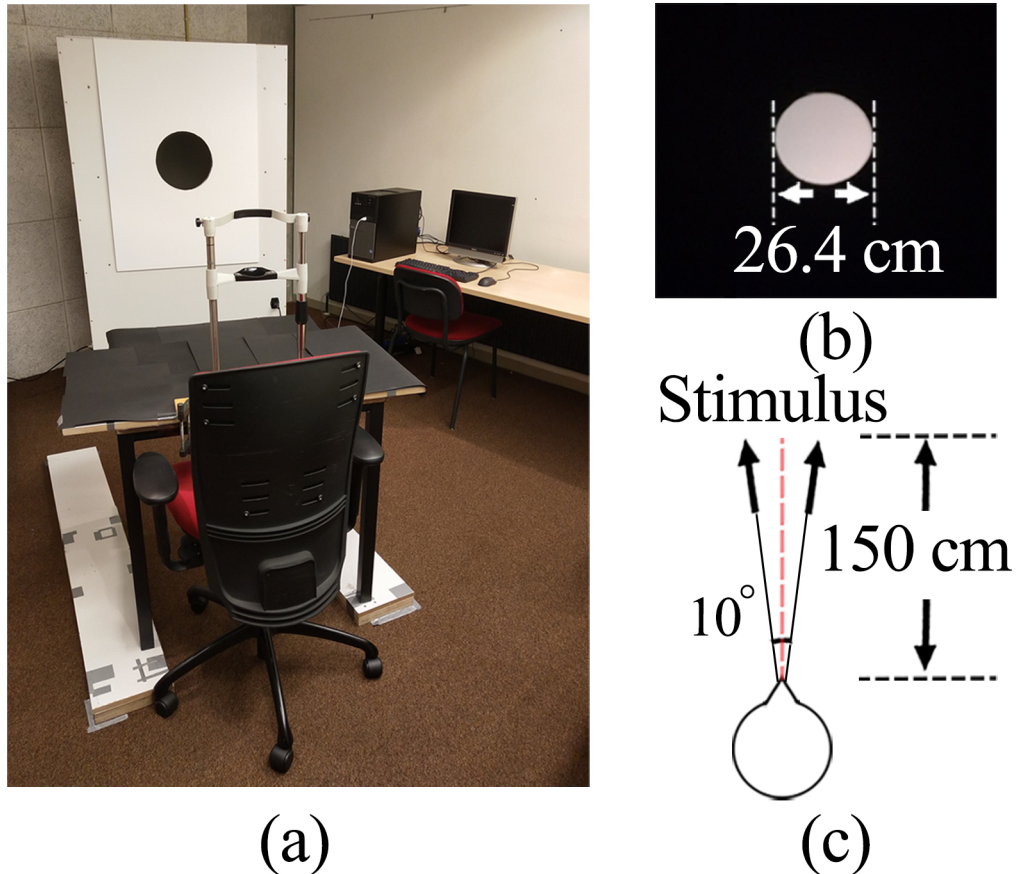


Figure 6.1: (a) Overview of the lab environment (b) Front view of the stimulus (c) Top view of the participant and stimulus.

colors (BC_8 and BC_9), which were less saturated versions of BC_2 and BC_3 , were added. The chromaticity coordinates of the base colors are shown in Table 6.1.

Four modulation directions were chosen, namely 0° (i.e. parallel to the direction of the u' axis), 45° , 90° (i.e. parallel to the v' axis) and 135° (as shown in Figure 6.2). The amplitude of the modulation varied between $0.00004 \Delta(u', v')$ and $0.05 \Delta(u', v')$. The luminance ratio of the two extreme colors of each sinusoidal modulation was based on the individual isoluminance, as determined by the preparation experiment. The average luminance of the stimulus was 35 cd/m^2 . The temporal frequency of the light modulation was 2, 4, 8, 10, 15, 20 or 25 Hz. Table 6.1 shows the variables of the 252 (i.e., 9 base colors \times 4 directions \times 7 frequencies) conditions.

6.2.3. PARTICIPANTS

Three participants performed the experiment (**AM**, female, 25 years; **MB**, male, 27 years; **XK**, male, 27 years). The participants received elaborate training and two of them had previous experience with chromatic flicker experiments. All participants had normal color vision, as measured with the Ishihara test for color deficiency and one of them was wearing glasses for corrected visual acuity. None of them were susceptible to migraine and/or epileptic seizures.

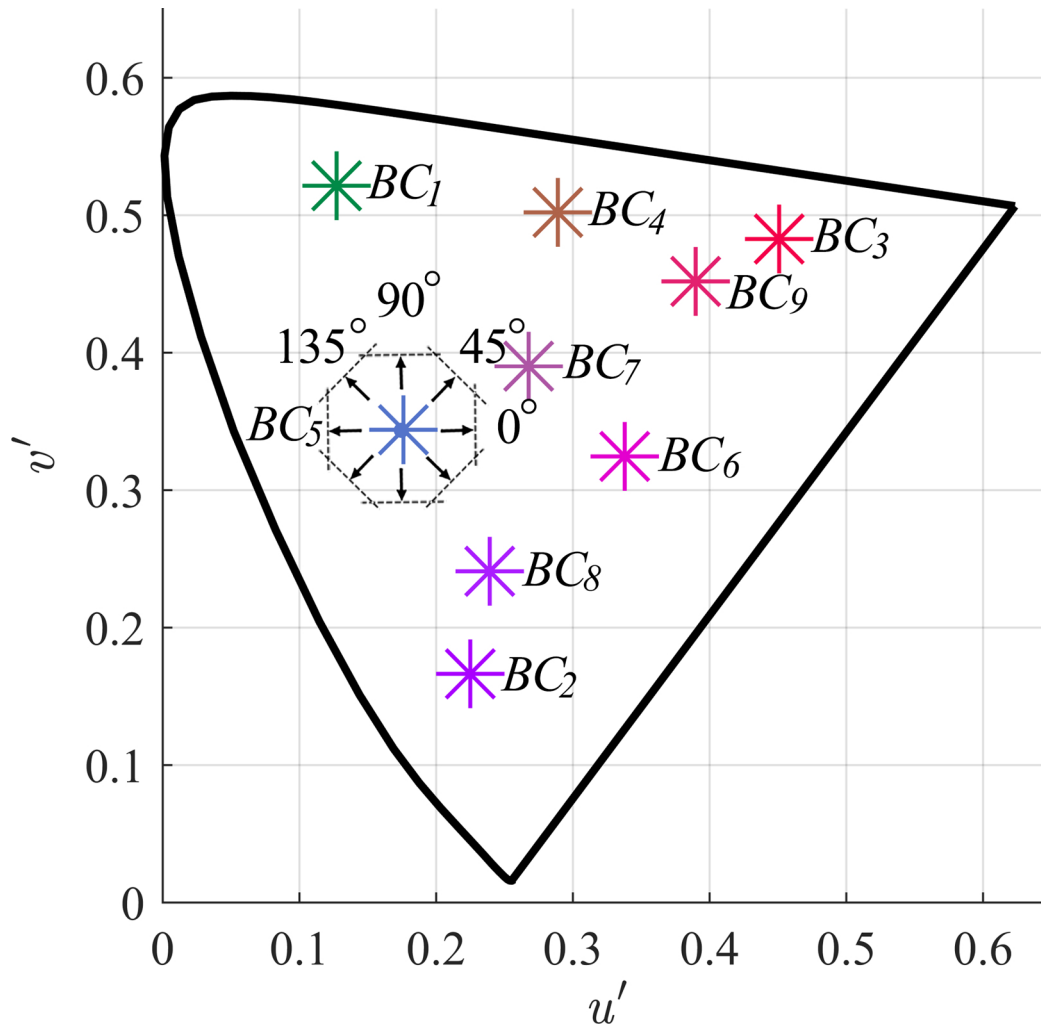


Figure 6.2: Nine base colors and four modulation directions (shown around BC_5) in the CIE 1976 UCS (u' , v') chromaticity diagram.

6.2.4. PROCEDURE

The experimental procedure was approved by the Daily Board of the Human-Technology Interaction group, Eindhoven University of Technology for ethical considerations. The main experiment was divided into different sessions on different days since it was too fatiguing for the participant to measure all conditions at once. The task of the experiment was explained both in text and orally by the experimenter. Before the start of the first session, participants were asked to sign an informed consent form, in which they were informed on their voluntary participation. Then an Ishihara color deficiency test was performed, followed by the collection of demographic information.

In each session, participants were instructed to sit down in the chair at the predefined position. They could adjust the chinrest and the height of the chair to a comfortable position before the experiment started. The method of adjustment was used to find the detection threshold of chromatic flicker. In **Chapter 5**, this method was shown to be both accurate and efficient to measure chromatic flicker thresholds. In order to correct for a possible error of anticipation, the adjustment was performed once with a relatively high starting amplitude and once with a low starting amplitude. In addition, participants were trained to

Table 6.1: The stimulus variables of the 252 conditions

BC	u'	v'	Modulation Direction	Frequency
1	0.1273	0.5213		
2	0.2251	0.1663		2 Hz
3	0.4509	0.4826		4 Hz
4	0.2891	0.5019	0°	8 Hz
5	0.1762	0.3438	45°	10 Hz
6	0.338	0.3245	90°	15 Hz
7	0.2678	0.3901	135°	20 Hz
8	0.2393	0.2409		25 Hz
9	0.3899	0.4517		

use a consistent decision criterion to minimize the variance within participants [104].

Each trial started with a beep sound, indicating the participant could start with the adjustment. The first stimulus of the trial had either a modulation amplitude of 0.05 $\Delta(u', v')$, for which flicker was clearly visible, or an amplitude of 0.0004 $\Delta(u', v')$, which appeared to be static. Participants were instructed to look at a fixation point in the center of the circular opening of the apparatus and to find the smallest amplitude at which the flicker was just visible, by increasing or decreasing the amplitude of the chromatic modulation. They could use the up and down arrow keys of a keyboard to increase or decrease the modulation amplitude with a large factor, while the left arrow and right arrow keys could be used to change the modulation amplitude with a small factor. The modulation amplitude was changed by a factor F :

$$F = \frac{A_{n+1}}{A_n} = 1.1^{\pm\alpha} \quad (6.1)$$

where A_n refers to the amplitude of the current stimulus, A_{n+1} refers to the amplitude of the next stimulus, and α equals 2 for the small factor, while α equals 5 for the large factor. When participants reached the lowest or highest amplitude of the stimulus range, they heard a warning signal indicating that they could not go lower or higher, respectively. When participants were satisfied with the stimulus that represented their detection threshold, they were instructed to look at the final stimulus for one or two seconds before deciding to really end the adjustment procedure. After pressing the Enter key, the next trial was presented.

The first session was a training session, which aimed to get the participants familiar with the method of adjustment and the experimental task. Participants were instructed to find a useful strategy and to stick to that strategy during the rest of the experiment. Subsequently, each base color was presented in a separate session, and the order of the base colors was the same for all participants (from BC_1 to BC_9). At the beginning of a session, a static adaptation stimulus with the chromaticity of the base color and a luminance of 37.5 cd/m^2 was presented for two minutes. Next, the conditions were presented in a random order. For each base color, there were 56 conditions (i.e., 7 frequencies \times 4 directions \times 2 starting amplitudes). Each session took about 30 minutes, and the experiment took 4.5 hours in total (without the preparation experiment and the training session).

After each session, participants were asked to write down some notes regarding their experience with the experiment. During the experiment, the chromaticity values of all stimuli that were shown were stored together with a time log. These data were used to look

at the strategy of the participant and for further analyses.

6.3. ANALYSES AND RESULTS

The experiment resulted in 504 detection thresholds (i.e. 252 conditions \times 2 starting amplitudes) expressed in CIE 1976 UCS (u' , v') for each participant. The thresholds can be plotted as a function of frequency for each base color, modulation direction and participant. Figure 6.3 shows an example of the data of participant **MB** at BC_1 for the four modulation directions. From Figure 6.3, we can see that the threshold increases as the frequency increases. There also seems to be an effect of modulation direction, which is most visible at 25 Hz. Moreover, the thresholds of the two starting amplitudes are quite consistent with each other.

These conclusions were confirmed by a linear mixed model analysis (LMM) with Detection Threshold as dependent variable and Base Color, Modulation Direction, Frequency and Starting Amplitude as fixed independent variables and with a random intercept for Participant. The analysis revealed a significant main effect of Base Color ($F(8, 1509) = 36.854$, $p < 0.001$), Modulation Direction ($F(3, 1509) = 126.435$, $p < 0.001$), Frequency ($F(6, 1509) = 892.323$, $p < 0.001$) and Starting Amplitude ($F(1, 1509) = 5.234$, $p = 0.022$). A low starting amplitude resulted in a slightly lower detection threshold compared to a high starting amplitude (MD = 0.001, $p = 0.022$), which was also found in Chapter 5. The detection thresholds were averaged over the two starting amplitudes for further analysis.

6.3.1. MODELING OF TCSFs

In order to model the TCSF for chromatic modulations, we expressed the data in terms of contrast sensitivity (i.e. the reciprocal of the detection threshold expressed as chromatic contrast). Three measures of contrast were used: $\Delta(u', v')$, ΔLMS (using Equation 6.2) and Δlms (using Equation 6.3).

$$\Delta LMS = \sqrt{(\Delta L)^2 + (\Delta M)^2 + (\Delta S)^2} \quad (6.2)$$

$$\Delta lms = \sqrt{(\Delta l)^2 + (\Delta m)^2 + (\Delta s)^2} \quad (6.3)$$

where $\Delta L, \Delta M, \Delta S$ are the differences between the L -, M -, and S -cone responses of the two extreme colors of the chromatic flicker stimulus at threshold modulation, which are calculated using the cone fundamentals of Stockman and Sharpe [133], and $\Delta l, \Delta m, \Delta s$ are the differences between the L -, M - and S -cone responses normalized with respect to the sum of the cone responses.

For each contrast measure, the contrast sensitivity was plotted as a function of frequency, resulting in 36 (i.e. 9 base colors \times 4 directions) TCSFs per participant. These TCSFs were fitted with:

$$S = \frac{1}{C} = e^{\beta_1 f + \beta_0} \quad (6.4)$$

where S is the contrast sensitivity, C the detection threshold and f the temporal frequency. Equation 6.4 can be rewritten as:

$$\ln S = \beta_1 f + \beta_0 \quad (6.5)$$

with β_1 being the slope and β_0 the intercept of the function.

For all three contrast measures, the average goodness-of-fit expressed in R^2 was higher than 0.93 and they were all very similar to each other (i.e., the average R^2 differed by a

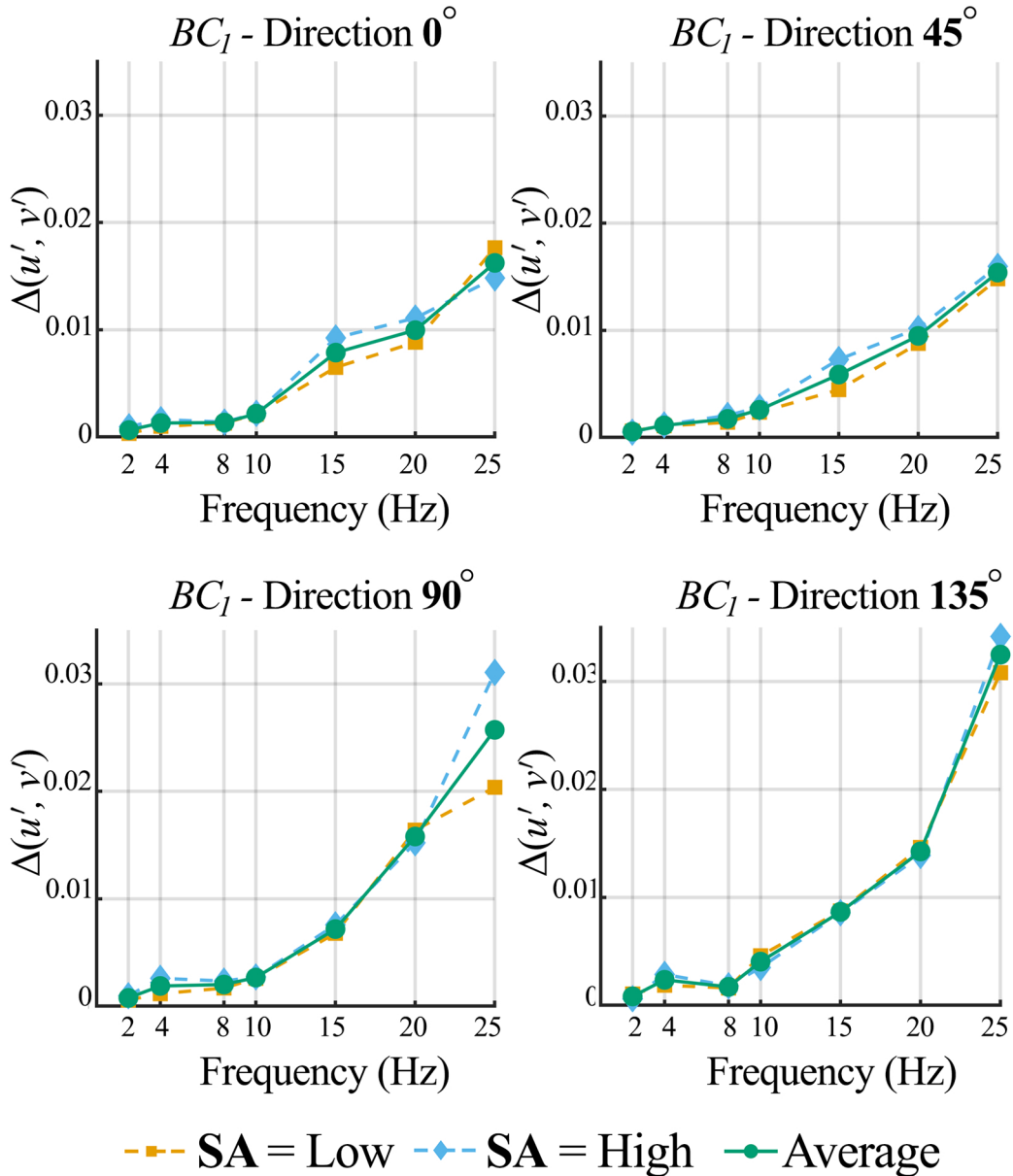


Figure 6.3: The chromatic flicker threshold in $\Delta(u', v')$ as a function of modulation frequency for BC_1 for the four modulation directions for participant **MB**. The two dashed lines show the thresholds for the two starting amplitudes (SA) and the solid line represents their average.

maximum of 0.0002). For further analysis, we chose ΔLMS as our contrast measure. The R^2 averaged over all conditions was 0.86, 0.97 and 0.96 for participant **AM**, **MB** and **XK**, respectively. As an example, Figure 6.4 shows the data and the fitted functions for all the TCSFs of participant **MB**.

Figure 6.5 gives the slopes (β_1 from Equation 6.5) and intercepts (β_0 from Equation 6.5) of the fits for the three participants. The figure shows that there are differences between the participants in the absolute values of the slope and intercept of the TCSF, but also in the effect of base colors and modulation direction, especially for the slope. In order to test the overall effect of base color and modulation direction, two separate linear mixed model (LMM) analyses were performed: (1) with Slope (β_1) as the dependent variable and (2) with

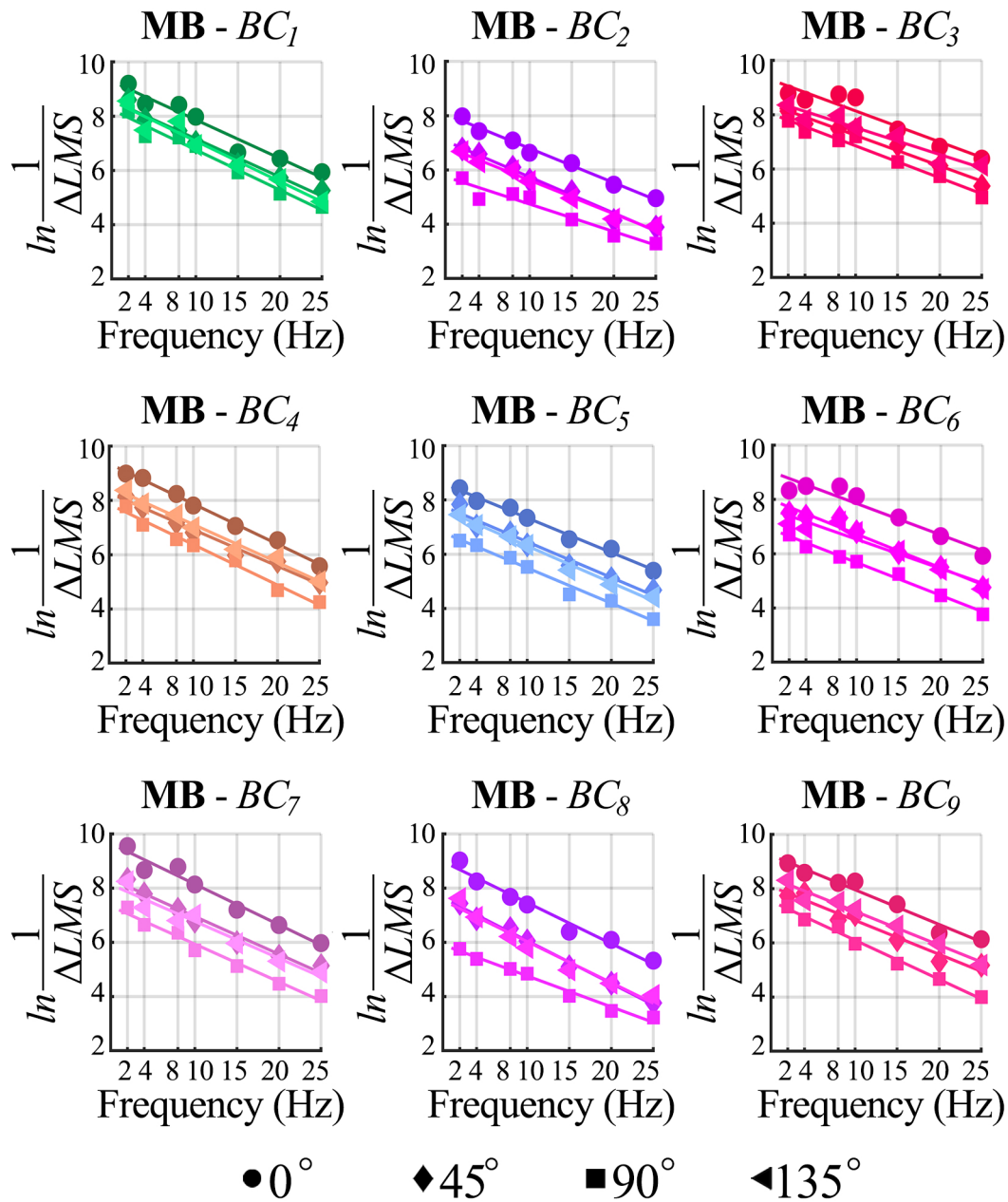


Figure 6.4: The contrast sensitivity as a function of frequency for all base colors (BC_1 to BC_9) and the four directions (0° , 45° , 90° and 135°) for participant **MB**. The straight lines represent the fitted functions according to Equation 6.5.

Intercept (β_0) as the dependent variable. In both analyses, Base Color and Modulation Direction were the fixed independent variables. We also included the interaction term between these two variables and a random intercept for Participant. The resulting p-values were obtained from a Type III sum of squares. Post-hoc analyses with Bonferroni correction were performed for the significant effects.

For the *Slope*, the LMM analysis showed a significant effect of *Base Color* ($F(8, 105) = 16.375$, $p < 0.001$). BC_1 had the lowest slope, BC_3 , BC_6 and BC_9 had the highest slope and the slopes of the other base colors were in between and not significantly different from each

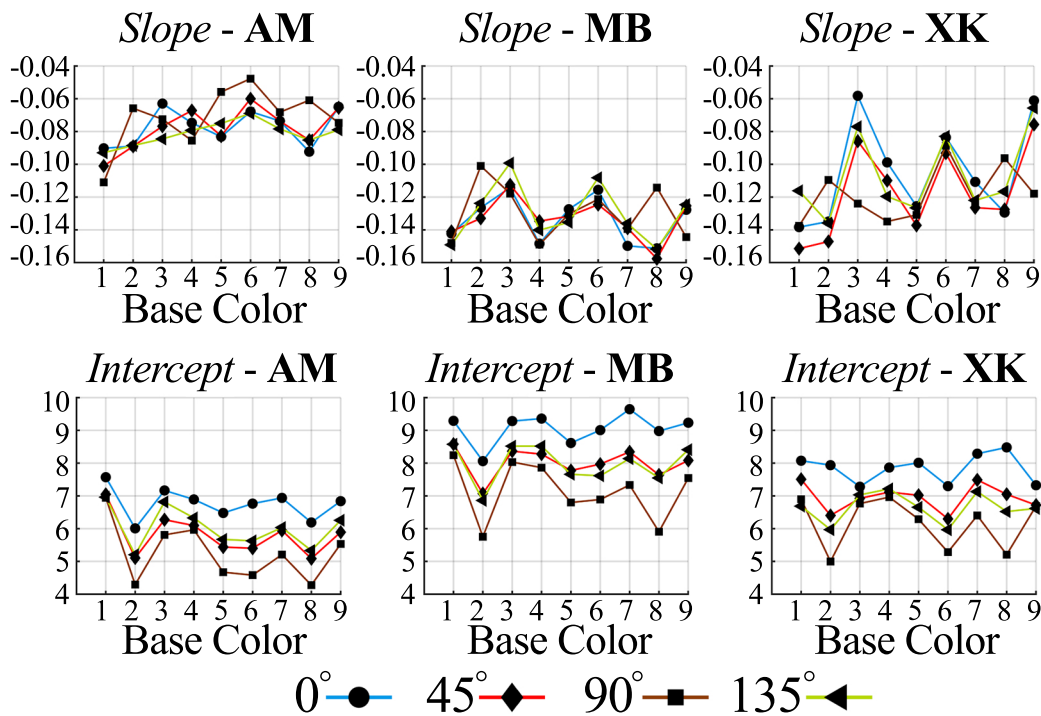


Figure 6.5: The slopes and intercepts of the fits of all conditions for the three participants.

other. There was also a significant interaction effect between *Base Color* and *Modulation Direction* ($F(24, 105) = 2.586, p < 0.001$). However, no significant main effect was found for *Modulation Direction* ($F(3, 105) = 0.762, p = 0.518$). The effect of *Modulation Direction* on *Slope* was different for different base colors, for example, BC_3 had the highest value of the slope at Direction 0° , while BC_2 had the highest value at Direction 90° .

For the *Intercept*, the effect of *Base Color* ($F(8, 105) = 41.949, p < 0.001$), *Modulation Direction* ($F(3, 105) = 179.595, p < 0.001$) and the interaction between *Base Color* and *Modulation Direction* ($F(24, 105) = 3.249, p < 0.001$) were significant. BC_2 had the lowest value of the intercept, which was significantly lower than all the other base colors, and BC_1 , BC_3 and BC_4 had the highest intercept. We also found that the intercept was lowest for Direction 90° and highest for Direction 0° . The intercept of the other two directions was in between and not significantly different from each other. For all base colors, the trend of the Intercept over the four modulation directions was the same. Direction 0° always had the highest Intercept, followed by Direction 135° , Direction 45° and Direction 90° . However, the difference was not significant for all base colors.

6.4. DISCUSSION

In this study, we measured the detection threshold of chromatic flicker and obtained TCSFs for nine base colors and four modulation directions for three participants. Instead of exploring inter-observer differences or modeling an average observer, we aimed at measuring and modeling individual observers. We found that the TCSF of all participants could be described by the same exponential model, the parameters of which changed from person to person. More research is needed to investigate if the model depends on demographic factors, such as the age of the observer.

The contrast sensitivity was defined as the reciprocal of ΔLMS , i.e. the difference between the LMS cone responses of the two extreme colors of the chromatic flicker stimulus at threshold modulation. The TCSFs were fitted with an exponential function (see Equation 6.4). The model fitted the data very well, with R^2 values around 0.90. The two parameters of the model (i.e. the slope β_1 and intercept β_0) were found to be dependent on the base color and modulation direction. This means that the visibility of a chromatic temporal modulation at a given frequency does not only depend on the ΔLMS of the two (extreme) colors of the modulation, but also on the LMS values themselves. So, ΔLMS itself cannot be used to accurately predict the visibility of temporal changes in light.

The results indicate that the chromatic TCSF is a low-pass function, which is in line with other studies using isoluminant stimuli [45, 54]. Even though we did not directly measure the CFF, we can conclude that it must be higher than 25 Hz, since participants still could see flicker at some modulation amplitudes. The CFF probably depends on base color and modulation direction, as can be seen in Figure 6.4.

In our experiment the average luminance of the chromatic flicker stimuli was fixed at 37.5 cd/m^2 . However, retinal illuminance has been found to influence the chromatic TCSF [136, 159]. Therefore, it is important to explore the effect of luminance level in future work.

Since the contrast sensitivity for chromatic temporal modulations can accurately be described by a simple exponential model for a wide range of color stimuli, it is not necessary to use a large number of frequencies to determine the TCSF in a certain location of the color space. Instead, only a few frequencies are enough, which would greatly reduce the duration of the experiment. For future studies, we plan to measure the TCSF for more base colors, modulation directions and average luminance values in order to ultimately develop a representative temporal color model.

6.5. CONCLUSION

The contrast sensitivity for chromatic temporal modulations expressed in $1/\Delta LMS$ was found to be a good indicator to describe the relationship with frequency. An exponential model with two parameters (i.e. slope and intercept) described the TCSFs with very high accuracy (with an average R^2 higher than 0.93 for the three observers). The two parameters of the model were found to significantly depend on the base color and direction of the chromatic modulation.



Modeling Sensitivity to Chromatic Temporal Modulations Using Individual Cone Fundamentals

Research has shown that spatial color models cannot accurately predict the perception of temporally modulated colored light and may lead to unpleasant observations such as jerkiness and flicker. Moreover, it is known that color perception differs between individuals, while most models are based on a standard observer. In this chapter, we collected data for the purpose of developing a temporal uniform color space taking into account individual differences in cone spectral sensitivities. Two psychophysical experiments were performed. In Experiment 1, we measured for fifteen participants the ratio in luminance between the two extreme colors of 60 chromatic flicker modulations at which the perceived flicker of that modulation was minimal. These so-called isoluminance ratios were used to estimate individual cone spectral sensitivities based on an adapted version of the Individual Colorimetric Observer (ICO) model of Asano, Fairchild, and Blondé [5]. The ICO model described the perceived luminance better than the cone spectral sensitivities of the standard observer. In Experiment 2, we measured the temporal contrast sensitivity at fifteen base colors, four modulation directions and three temporal frequencies for three participants that had substantially different isoluminance ratios. The results give a first impression on the size of individual differences in sensitivity to chromatic temporal modulations. In addition, we propose two quantitative measures, i.e., *circularity* and *homogeneity*, to describe the local and global uniformity of color spaces respectively. In general, the DKL space is somewhat more uniform than the LMS space and the ICO-based cone fundamentals improve the temporal uniformity of the color spaces both locally and globally. However, contrary to our expectations, the $u'_{10^\circ} v'_{10^\circ}$ color space has overall the best circularity and homogeneity. We present a number of possible explanations and suggestions for further research to design a (more) perceptually uniform temporal color space.

This chapter is submitted to Lighting Research & Technology with slight adaptations as: "Modeling sensitivity to chromatic temporal modulations using individual cone fundamentals", Xiangzhen Kong, Ingrid Vogels,

7.1. INTRODUCTION

7.1.1. BACKGROUND - THE NEED FOR A TEMPORALLY UNIFORM COLOR SPACE

Light emitting diodes (LEDs) have enabled easy and inexpensive ways to create dynamic colored light. However, it is not straightforward to create pleasant and high-quality dynamic light effects, as guidelines for temporal color differences and smooth color transitions are still missing. As a result, dynamic light might result in unpleasant artifacts, such as jerkiness and flicker.

The perception of *spatial* color differences is much better understood. Spatial color vision is described in established color spaces. Since the development of the Munsell color space in the early 20th century, several color spaces have been developed to describe spatial color differences in a perceptually uniform way (e.g., CIE 1976 UCS and CIELAB). A perceptually uniform *spatial* color space ensures that the perceived difference between two spatially separated colors is proportional to the distance (usually Euclidian) between the colors, irrespective of the location in the color space. The perceived difference between two colors can then be expressed in a color difference metric, such as ΔE_{ab}^* in CIELAB, for an average (or standard) observer.

Since spatial properties of a visual stimulus are processed in different areas of the visual cortex than temporal properties [57], it is unlikely that ΔE_{ab}^* can accurately predict temporal color differences as well. Several studies have shown that ΔE_{ab}^* is indeed not an accurate measure to describe a temporal change in color. For example, Sekulovski et al. [121] reported that the visibility threshold of smoothness, i.e., the maximum color difference between two successive colors that is allowed in order to perceive a temporal color transition as smooth, depended on the direction of the color change in CIELAB. More specifically, the threshold was about ten times smaller for temporal changes in lightness than for chroma or hue. In **Chapter 3** and **Chapter 4**, we found that two color transitions with the same color change per second in CIELAB (i.e., $\Delta E_{ab}^*/s$) were not necessarily perceived as having the same speed. The actual speeds in $\Delta E_{ab}^*/s$ of the color pairs tested had to differ by a factor ranging between 0.56 and 1.49 in order to be perceived as equal.

These studies illustrate the urgency for a perceptually uniform temporal color space to create appealing dynamic light effects. Such a temporal color space could be developed based on data describing the human sensitivity to periodic temporal modulations in luminance and/or chromaticity, both depending on the modulation frequency [71]. To be more specific, a perceptually uniform temporal color space should have the following properties. First, for any given color point, the temporal modulation that is just visible (called detection threshold) should be independent of the modulation direction in the color space. This property describes the local uniformity of the color space and is referred to as *circularity*. Second, for any given modulation direction, the detection threshold should be independent of the color point. This property is referred to as *homogeneity* and describes the global uniformity of the color space. Third, these requirements should hold at any temporal frequency. The thresholds themselves may depend on frequency, but the effect of frequency should be the same for all color points and modulation directions. This property is called *frequency-consistency*. Besides these uniformity requirements, the perceptually uniform temporal color space should take individual differences into consideration in order to be useful for practical implementations. As such, people may have different sensitivities to

chromatic modulations, as long as the same color difference metric can be used to describe the perception of these chromatic modulations. To the best of our knowledge, the circularity and homogeneity of temporal color spaces have not been evaluated systematically. Among the limited studies, Safdar et al. [112] proposed STRESS-based [29] metrics to quantify the local and global uniformity for color discrimination ellipses. These metrics are vector based, and start from fitting ellipses through visibility thresholds. Since this is less appropriate and accurate for our application, we have used alternative metrics in our evaluation.

The relation between temporal frequency and sensitivity to temporal modulations is well-known and described in literature by so-called temporal contrast sensitivity functions (TCSFs). The TCSF for luminance modulations has been measured and modeled extensively (e.g. [13, 159]). For chromatic modulations, Dobkins, Anderson, and Lia [15] and **Chapter 6** showed that a single exponential function can be used to model the TCSF. Furthermore, in **Chapter 6**, we found that the parameters of the function depended on the location in the color space (both in CIE 1976 UCS (u' , v') and in the *LMS* cone excitation space), and the parameters varied between the three individuals used in this study. In order to develop a perceptually uniform temporal color space, it is important to understand the differences between individuals. Therefore, in this study, we aim to measure the variability among individuals and its effect on the uniformity of existing color spaces.

7.1.2. INDIVIDUAL DIFFERENCES

Properties of spatial vision are often described for a standard observer. This standard observer is usually determined as a group average of empirical data and transformed into a mathematical model. The concept of a standard observer has been used to describe various properties of human vision, such as luminance (i.e., the standard photopic luminous observer, characterized by the luminous efficiency function), chromaticity (i.e., the standard colorimetric observer, characterized by Color Matching Functions, CMFs) and spatial contrast (i.e., the spatial standard observer, defined by Watson [149]). CMFs are widely used as they can be linearly transformed from one set to another set of primaries, such as the L-, M- and S-cone primaries [133]. The latter primaries refer to the cone spectral sensitivities and are also sometimes referred to as cone fundamentals [133, 143]. It would be valuable to define a standard temporal observer for describing perceived temporal contrast. This standard observer could then be used to develop a temporal color space. It is, however, important to realize that a standard observer is different from any real observer [69], since individuals can differ in their perception of color [80].

Individual differences in visual perception can have several causes. From a physiological point of view, individual differences in vision can be explained by differences in pre-receptoral screening (i.e., related to the density of the lens and macular pigment), differences in photopigment and cone ratios, and post-receptoral differences [94, 154]. In addition, aging is an important factor causing physiological changes in the visual system [92]. Most of the research on individual differences in vision has focused on understanding perceptual processes from a more fundamental perspective, and hence try to explain where individual differences originate from. Though very relevant and of great value in itself, it is insufficient to understand how important these differences are in real applications.

Literature has shown that individual differences in the temporal contrast sensitivity function can be substantial. For example, Dobkins, Lia, and Teller [16] and Dobkins, Anderson, and Lia [15] compared the TCSFs between infants and adults, and inferred from

that insights in the development of the human visual system. In addition, in **Chapter 6**, we found significant differences in the TCSFs between individuals, not as much in the specific function used to model the data, but rather in the values of the model parameters. This suggests that it is important to investigate these individual differences in more detail and to consider the consequences for designing temporal color transitions that are meant to be pleasant for all observers.

7.1.3. ISOLUMINANCE

Luminance information and chromaticity information are processed in different channels of the human visual system [73]. Hence, for measuring the sensitivity to chromatic temporal contrast it is essential to only activate the chromatic channels by using isoluminant stimuli. This can be done for an average observer using the standard photopic luminous efficiency function $V(\lambda)$. However, as mentioned above, this function does not represent the sensitivity of an individual observer [52]. In other words, even though the physical luminance of a temporal stimulus is constant over time, observers may perceive the luminance to change due to individual deviations from the luminous efficiency function of the standard observer [82].

According to CIE, luminance is defined as:

$$L = K_m \int L_{e,\lambda} V(\lambda) d\lambda \quad (7.1)$$

where L is the luminance in units of cd/m^2 ; K_m is the maximum luminous efficiency, which is defined to be $683 \text{ lm}/\text{W}$; $L_{e,\lambda}$ is the spectral radiance of the stimulus; and $V(\lambda)$ is the CIE photopic luminous efficiency function. The sensation luminance (SL) is almost identical to Equation 7.1 except for the critical difference in capitalization, with $\nu(\lambda)$ being the individual luminous efficiency function:

$$SL = K_m \int L_{e,\lambda} \nu(\lambda) d\lambda \quad (7.2)$$

Given the importance of differences in $\nu(\lambda)$, isoluminant stimuli have to be defined per individual before measuring their chromatic TCSFs [93]. Several techniques have been used to configure individual isoluminance, such as minimum motion technique, minimum-border or minimally distinct-border method, (MDB), heterochromatic brightness matching and heterochromatic flicker photometry (HFP). For reasons of practical convenience, researchers often employ only one method. However, the results may differ largely between different methods, as found by Wagner and Boynton [147] who compared four methods of heterochromatic photometry. The detailed comparison of these methods is out of scope of this study. Since heterochromatic flicker photometry is more recently considered as a reliable and accurate method to measure a subject's luminous efficiency function [72], we employ this method to configure individual isoluminance in this study.

7.1.4. CONTRAST SENSITIVITY FOR CHROMATIC TEMPORAL MODULATIONS

Contrast sensitivity for chromatic temporal modulations has been studied in literature using different types of stimuli, among which spatiotemporal stimuli are the most common (see [13]). For example, Dobkins, Anderson, and Lia [15] used spatiotemporal stimuli like moving horizontally-oriented sinusoidal red/green gratings subtending a $15 \times 15^\circ$ visual angle presented at different temporal frequencies, which corresponded to different speeds

expressed in units of deg/s. Kelly and Norren [55] used a 1.8° uniform stimulus of which the chromaticity was modulated in time by changing the relative phase of a red and green sinusoidal waveform. In **Chapter 6**, we presented a 10° uniform field of which the chromaticity was temporally modulated in different modulation directions by changing the relative contribution of the red, green and blue primary.

The first step of modeling chromatic TCSF is to find a suitable measure to express the sensitivity to chromatic temporal modulations, independent of the type of stimulus that is used. Since chromaticity is a two-dimensional property, it is not straightforward to do so. Luminance modulations are one-dimensional and can easily be expressed in terms of modulation amplitude or Michelson Contrast [111]. The sensitivity to red-green chromatic modulations has been expressed in terms of the luminance contrast of either the red or green signal [55]. More physiology-based measures, as the contrast of the long-wavelength sensitive (*L*-), medium-wavelength sensitive (*M*-), and short-wavelength sensitive (*S*-) cone responses, have been employed as well [89]. The TCSFs for *L*-, *M*-, and *S*-cones have been measured separately [47, 134], but due to post-receptoral and cortical processing, it is far from intuitive to derive from them the resulting perception for any specific modulation. As such, the physiologically based models mainly predict the perception of stimuli that activate only one channel of the visual system, e.g., the red-green opponent channel [20]. In the absence of a better alternative, we previously used the sum of squared *L*-, *M*-, and *S*-cone response differences to describe the sensitivity to chromatic modulations in various directions in **Chapter 6**.

The next step is to model the TCSF. It is by now well known that the TCSF for luminance modulations under photopic vision is a band-pass function, where the peak sensitivity and corresponding frequency depend on the adapting luminance [136]. For chromatic modulations, the TCSF is typically a low-pass function where the sensitivity decreases with increasing frequency [16, 44]. This function can be modeled as a single exponential function:

$$S = \frac{1}{C} = e^{\beta_1 f + \beta_0} \quad (7.3)$$

where S is the contrast sensitivity, C is the contrast, f is the temporal frequency, β_0 is called the *intercept* and β_1 is called the *slope*. The function has proven its validity in literature [16, 60]. Despite of the simplicity of its form, it is important to emphasize that TCSF is not a single invariant function, but that its parameters are dependent on many factors, such as the spatial configuration of the stimulus, its background intensity, duration of the stimulus, its eccentricity in the visual field, the use of eye movements, the psychophysical method used to determine the thresholds and the observer [60, 151]. The observer dependency is not surprising, but quantitative comparisons between individuals are lacking. In addition, due to technical limitations in the past, most experiments on chromatic TCSFs have been carried out for a limited set of color stimuli.

7.1.5. RESEARCH OBJECTIVES

The aim of this study is to collect more data on the sensitivity of individuals to chromatic temporal modulations and to quantify the uniformity of existing color spaces. To do so, we created light modulations around fifteen chromaticity points with four modulation directions to cover a large part of the color space. First, we measured the ratio in luminance between the two extreme colors of these chromatic flicker modulations at which the perceived flicker of that modulation was minimal, referred to as isoluminance ratio. From these

ratios we estimated the cone spectral sensitivities of fifteen participants and selected three participants with (relatively) large differences. Then, we measured the temporal chromatic sensitivity of these participants for the fifteen chromaticity points and four modulation directions. Next, we evaluated the performance of existing color spaces (i.e., LMS and DKL) in terms of *local uniformity* (i.e., *circularity*), *global uniformity* (i.e., *homogeneity*), and *frequency-consistency*, and tested if these models improve when individual cone spectral sensitivities are taken into account.

7.2. METHODS

Two experiments were carried out. In Experiment 1 we measured individual isoluminance ratios for a limited set of chromatic modulations with fifteen participants. In Experiment 2 we measured for three participants the detection thresholds of a broad range of isoluminant chromatic flicker stimuli. Both experiments were approved by the Daily Board of Human-Technology Interaction Group at Eindhoven University of Technology on ethical considerations. In this section we describe the experimental setup and the general characteristics of the stimuli and the procedure, as far as they were the same for both experiments. Specific information for each experiment is described in sections 3 and 4, respectively.

7.2.1. APPARATUS

The experimental setup consisted of a customized LED lighting system mounted in a box. The lighting system consisted of 36 Cree[®] XLamp[®] XP-E LEDs, including 12 Red LEDs (dominant wavelength: 620-630 nm), 8 Green LEDs (dominant wavelength: 520-535 nm) and 16 (Royal) Blue LEDs (dominant wavelength: 450-465 nm). The LEDs were driven by means of Pulse Width Modulation (PWM) technology at a driving frequency of 2,000 Hz. The PWM signals were generated by an Arduino Due microcontroller, for each channel with 16 bits. The lighting system was calibrated before the experiment with a spectrometer (JETI[®] specbos 1201), and it received *RGB* values as input.

The lightbox (width: 1.5 m, depth: 0.8 m, height: 1.5 m) was placed in front of the participants as shown in Figure 7.1(a). The inner surfaces of the box were painted with white chalk paint to obtain matte diffuse reflecting surfaces. The box had a circular opening of 26.4-cm diameter (Figure 7.1(b)). The 36 LEDs were placed inside the box such that the light reflected through the opening was quite uniform (i.e., the luminance in the circular area deviated by a maximum of 3.3% around the average level of 37.5 cd/m², while the chromaticity deviated by a maximum of $\Delta(u', v') = 0.0013$).

Participants were seated in a chair and looked at the light reflected through the circular opening. In front of them was a table with a keyboard, to enter their responses, and a chin rest, which was adopted to limit head movements and to make sure that their eye level was at the center of the circular opening (1.3 m above the floor). Participants were positioned 1.5 m from the front surface of the box, which resulted in a commonly used 10° visual field (Figure 7.1(c)).

7.2.2. GENERATING THE STIMULI

The stimuli were defined in the chromaticity diagram which is similar to CIE 1976 UCS (u', v') chromaticity diagram, but the coordinates were computed with CIE 10-degree observers

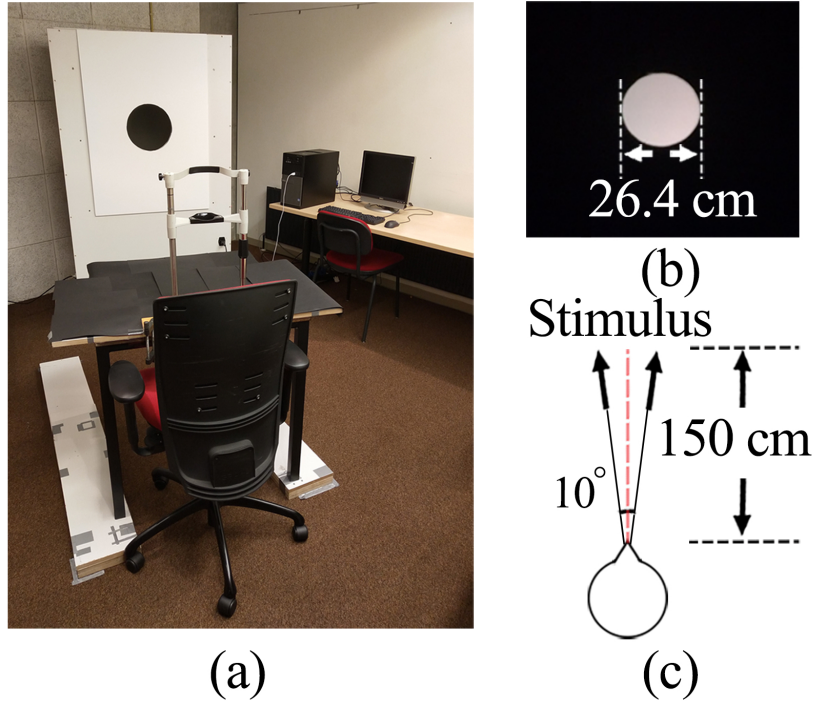


Figure 7.1: (a) Overview of the lab environment, (b) front view of the stimulus, and (c) top view of the observer and stimulus. During the experiment, the ambient was dark and the light reflected from the circular opening was the only light source.

instead of CIE 2-degree observers, using Equation 7.4.

$$\begin{cases} u'_{10^\circ} = \frac{4X_{10^\circ}}{X_{10^\circ} + 15Y_{10^\circ} + 3Z_{10^\circ}} \\ v'_{10^\circ} = \frac{9Y_{10^\circ}}{X_{10^\circ} + 15Y_{10^\circ} + 3Z_{10^\circ}} \end{cases} \quad (7.4)$$

where X_{10° , Y_{10° , and Z_{10° are the tristimulus values computed with the CIE 10-degree observers. They were transformed to *RGB* values to generate the input for the LEDs. The chromaticity of the stimuli varied sinusoidally around a base color BC with chromaticity coordinates $(u'_{10^\circ}, v'_{10^\circ})$, in a predefined modulation direction θ , with a modulation amplitude A expressed in $\Delta(u', v')$, and a temporal frequency F ($= 1/T$, T is the period of the sinusoidal amplitude modulation) (see Figure 7.2(a)). The amplitude A is defined as the distance between the two extreme colors C_1 and C_2 , which can be calculated using the following equations:

$$\begin{cases} C_i(u'_{10^\circ}) = u'_{10^\circ BC} + (-1)^i \frac{A}{2} \cos \theta \\ C_i(v'_{10^\circ}) = v'_{10^\circ BC} + (-1)^i \frac{A}{2} \sin \theta \end{cases} \quad (i = 1, 2) \quad (7.5)$$

The chromaticity of the stimulus changed along one of four directions, as shown in Figure 7.2(b), namely 0° (i.e., parallel to the direction of the u'_{10° axis), 45° , 90° (i.e., parallel to the v'_{10° axis) and 135° . The luminance of the stimuli was not constant. In Experiment 1, the luminance ratio between the extreme colors C_1 and C_2 varied in order to measure

Table 7.1: The chromaticity coordinates of the fifteen base colors in chromaticity diagram

BC	u'_{10°	v'_{10°	BC	u'_{10°	v'_{10°	BC	u'_{10°	v'_{10°
1	0.1273	0.5213	6	0.2209	0.4338	11	0.2866	0.4164
2	0.1553	0.4512	7	0.2393	0.2409	12	0.3146	0.3463
3	0.1833	0.3811	8	0.2489	0.3637	13	0.3242	0.4691
4	0.1929	0.5039	9	0.2586	0.4865	14	0.3522	0.399
5	0.2113	0.311	10	0.2769	0.2936	15	0.3899	0.4517

isoluminance. In Experiment 2, the luminance of the stimulus was manipulated to be isoluminant for each individual participant.

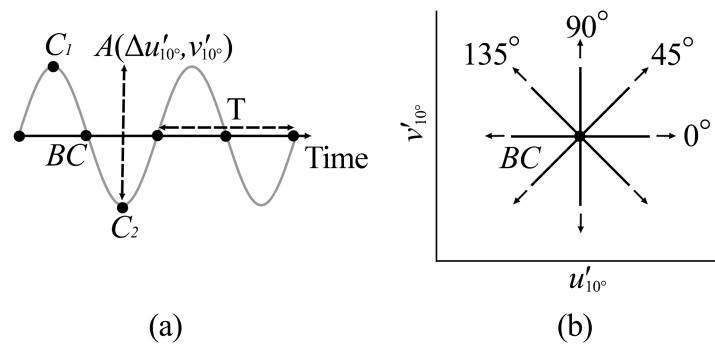


Figure 7.2: (a) Sinusoidal modulation between two extreme colors C_1 and C_2 around the base color BC, and for a modulation amplitude A expressed in $\Delta(u'_{10^\circ}, v'_{10^\circ})$. (b) The four directions of chromatic modulation used around a base color BC shown in the $u'_{10^\circ}, v'_{10^\circ}$ chromaticity diagram.

Fifteen (BC_1 to BC_{15}) evenly distributed color points in chromaticity diagram were chosen as base colors, as shown in Figure 7.3. The corresponding chromaticity coordinates are given in Table 7.1.

The choice for these base colors was partly based on previous experiments. We showed in Chapter 6 that participants couldn't reliably perform the task of Experiment 1 for highly saturated blue and red colors. Therefore, we limited ourselves to the less saturated base colors BC_7 and BC_{15} (which were also used in our previous experiments) that together with the base color BC_1 spanned the triangle of base colors used for the two experiments reported in this chapter.

7.2.3. GENERAL PROCEDURE

The participants were recruited in Eindhoven University of Technology. They were not allowed to participate if they had any medical history of epileptic seizures or migraine, or if they were very sensitive to flickering lights. Before the start of each experiment, the participants were asked to read and sign the informed consent form, in which they were informed of their voluntary participation. We then administered the Ishihara test for color deficiency, and asked the participants for their demographic information.

The task of each experiment was explained both in text and orally by the experimenter. The observers were instructed to sit down in the chair at the predefined position. The chin

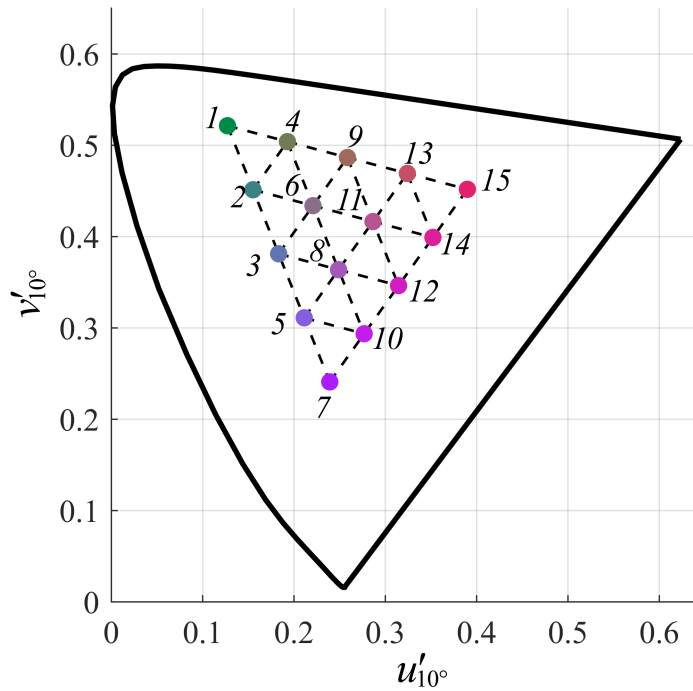


Figure 7.3: Fifteen base colors in $u'_{10^\circ} v'_{10^\circ}$ chromaticity diagram. The number assigned to a given BC was in the order of increasing u'_{10° -coordinate in $u'_{10^\circ} v'_{10^\circ}$.

rest and the height of the chair were adjusted such that the participant's eyes were at the same height as the center of the circular opening in the light box. We used the method of adjustment for both experiments. A standard keyboard with United States layout was provided as an input device for the observers. All the data (i.e., u'_{10° , v'_{10° , luminance, pressed keys) were stored together with the time log. These data were used for further analyses. Whenever necessary, the experimenter answered questions asked by the observers.

7.3. EXPERIMENT 1: ISOLUMINANCE

7.3.1. PARTICIPANTS

Fifteen people participated in this experiment. Five of them were female and ten were male; their age varied between 21 and 27 years ($M = 23.3$, $SD = 1.8$). Five of the observers were wearing contact lenses or glasses to correct their visual acuity. All participants had normal color vision according to the Ishihara test.

7.3.2. STIMULI

All stimuli were sinusoidal chromatic modulations around one of the fifteen base colors with a fixed amplitude A of $\Delta(u'_{10^\circ}, v'_{10^\circ}) = 0.05$ and a modulation frequency of 25 Hz, which is a commonly used frequency for HFP (e.g. [123]). Four modulation directions were used for each base color, resulting in 60 (i.e., 15 *Base Color* \times 4 *Modulation Direction*) conditions in total.

7.3.3. PROCEDURE

The task of the participants was to find the luminance ratio LR between the two extreme colors C_1 and C_2 of the modulation, at which flicker was perceived as minimal. They could increase the ratio with small steps by pressing the Right-key and with large steps by pressing the Up-key, and decrease the ratio with small steps using the Left-key and with large steps using the Down-key. The luminance ratio was changed according to the following equation:

$$\frac{LR_{n+1}}{LR_n} = 1.01^{\pm\alpha} \quad (7.6)$$

where LR_n refers to the current luminance ratio, LR_{n+1} to the luminance ratio of the next stimulus, and α was 1 for the fine tuning (i.e., a change of about 1%) and 5 for the rough tuning (i.e., a change of about 5%). The luminance ratio LR was defined as LC_2/LC_1 , and the average luminance of the stimulus was kept at a constant level of 37.5 cd/m^2 during the adjustment.

In order to reduce the error of anticipation, the initial value of LR was once 0.8195 (i.e., 1.01^{-20}) and the other time 1.2202 (i.e., 1.01^{20}). Both initial luminance ratios were expected to induce noticeable flicker for all participants. Thus, all participants assessed all 60 conditions twice. To keep all stimuli within the gamut of the system, the luminance ratio had to stay within the range between 0.8 and 1.5. Once the luminance ratio went beyond this range, a warning sound was played, indicating that LR could no longer be decreased or increased, respectively. When the participant was satisfied with the stimulus that appeared to have minimal flicker, (s)he was instructed to look at the final stimulus for one or two seconds before deciding to press the Enter-key to finish the trial. A beep sound was played to indicate the start of a new trial.

We estimated from a pilot study that Experiment 1 would last at least 1.5 hours per participant. The stimuli were divided into fifteen different blocks with a fixed base color, to keep the chromatic adaptation point as constant as possible. Within a block, all 8 stimuli (i.e., $4 \text{ Modulation Direction} \times 2 \text{ Initial Luminance Ratio}$) with the same base color were randomized. Before the start of a new base color block, a two-minute period (see Figure 7.4(a)) was included for the participant to adapt to the new base color at a luminance level of 37.5 cd/m^2 . Two minutes were found to be sufficient for a 100% of steady-state adaptation [22]. In addition, to minimize fatigue and eye-strain, we divided the blocks over two sessions (as shown in Figure 7.4(b)). Each session consisted of multiple blocks of seven or eight base colors with a proper break halfway. A 5-to-15-minute break was possible, depending on the participant's preference. The order of the fifteen base colors was counterbalanced between participants using a Latin-square design.

7.3.4. RESULTS

ISOLUMINANCE RATIOS

In Experiment 1, 120 luminance ratios (i.e., $15 \text{ Base Color} \times 4 \text{ Modulation Direction} \times 2 \text{ Initial Luminance Ratio}$) were collected per participant. The data were analyzed with an UNIANOVA analysis with *Luminance Ratio* as dependent variable, and *Base Color*, *Modulation Direction* and *Initial Luminance Ratio* as fixed independent factors, including their interaction terms. *Participant* was included as a random factor, and an intercept term was added to the model as well. The analysis was performed using IBM Statistics for Windows (Version 23).

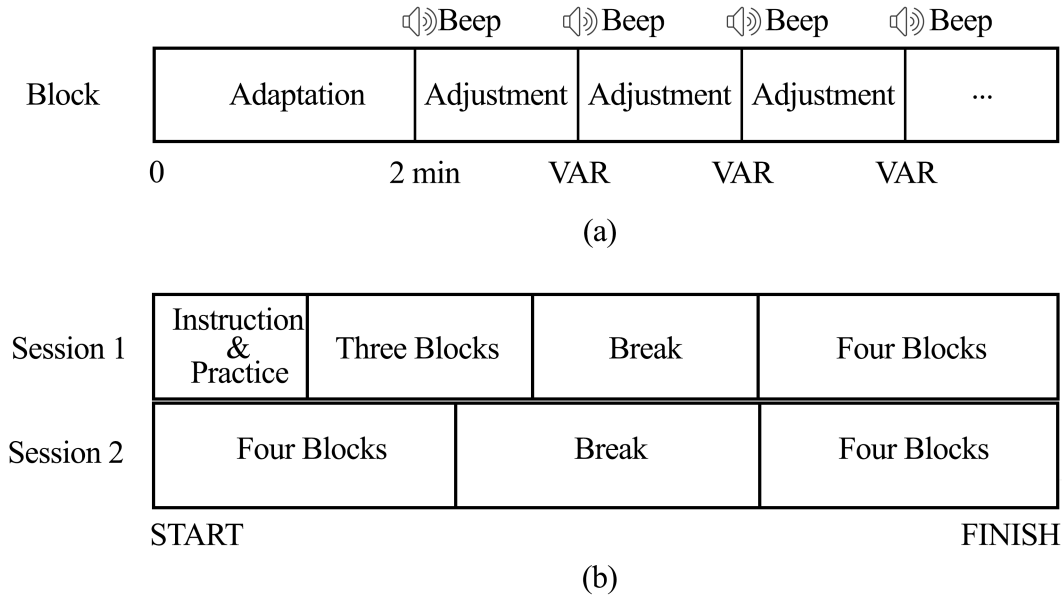


Figure 7.4: (a) Illustration of each base color block. Each adjustment trial started with a beep sound. VAR stands for variable adjustment time. (b) The timeline for the two sessions.

We found that the main effect of *Initial Luminance Ratio* was significant ($F(1, 1708) = 35.118, p < 0.001, \text{partial } \eta^2 = 0.020$). The mean luminance ratio (values around 1) was on average 0.008 smaller when the trial started with a low initial luminance ratio compared to a high initial luminance ratio. Other studies have also reported that the initial value of the method of adjustment might affect the results [50]. We averaged the results of the two initial luminance ratios to get the best estimation for the “true” luminance ratio at which the two stimuli are perceived as isoluminant. The resulting 60 isoluminance ratios (i.e., 15 *Base Color* \times 4 *Modulation Direction*) averaged over all fifteen participants are shown in Figure 7.5.

The analysis also revealed a statistically significant main effect of *Base Color* ($F(14, 1708) = 7.925, p < 0.001, \text{partial } \eta^2 = 0.061$) and *Modulation Direction* ($F(3, 1708) = 36.404, p < 0.001, \text{partial } \eta^2 = 0.060$) and a significant interaction between *Base Color* and *Modulation Direction* ($F(42, 1708) = 13.305, p < 0.001, \text{partial } \eta^2 = 0.247$). This indicates that the luminance ratio to minimize perceived flicker depends both on the position in the chromaticity diagram and on the direction of the modulation. We also found a significant main effect of *Participant* ($F(14, 1708) = 15.734, p < 0.001, \text{partial } \eta^2 = 0.114$), confirming the importance of individually adjusting the luminance ratio to minimize perceived flicker. The maximum difference in average luminance ratio between two participants was as large as 0.029.

ESTIMATING INDIVIDUAL CONE FUNDAMENTALS

The isoluminance ratios are not only useful to create isoluminant chromatic flicker stimuli, but they can also be used to estimate the individual relative spectral sensitivities of *L*-, *M*-, and *S*-cones. For a standard observer, the *L*-, *M*- and *S*-cone fundamentals (hereafter referred to as cone fundamentals) can be derived from the standard CMFs; for instance, the 10° cone fundamentals are based on the Stiles & Burch 10° CMFs [133]. As shown above, isoluminant ratios clearly differ between individuals, which indicates that the cone fundamentals are also different from person to person. In order to estimate individual cone

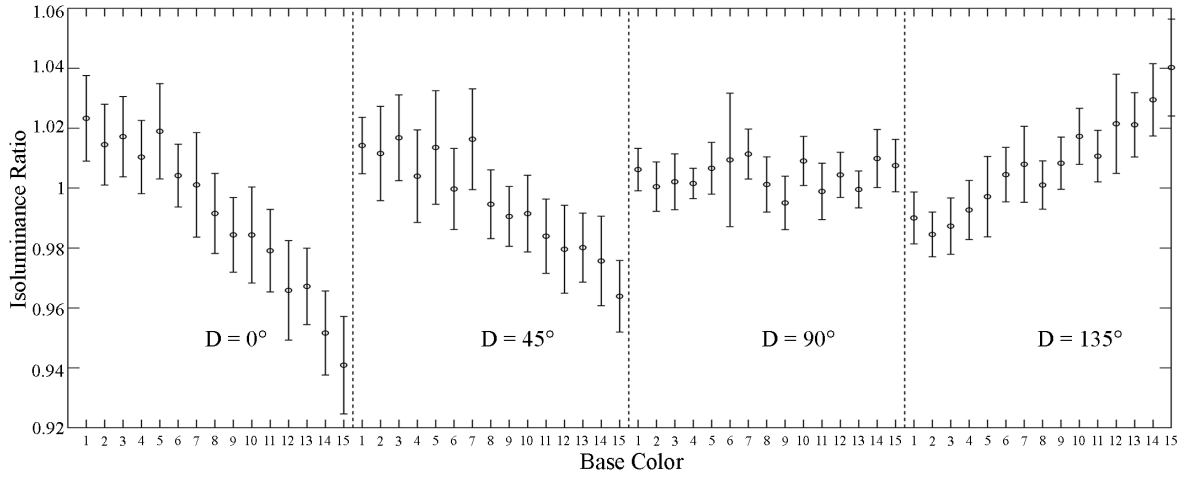


Figure 7.5: Plot of the isoluminance ratio per condition (i.e., *Base Color* × *Modulation Direction*) averaged over all fifteen participants and the two initial luminance ratios. The symbols represent the mean isoluminance ratio, and the error bars indicate the 95% confidence interval.

fundamentals, we used the individual colorimetric observer (ICO) model proposed by Asano, Fairchild, and Blondé [5], which is an extension to CIE06PO [69]. This ICO model can be expressed as follows:

$$lms_{ICO} = f(a, v, d_{lens}, d_{macula}, d_L, d_M, d_S, s_L, s_M, s_S) \quad (7.7)$$

where lms_{ICO} is an abbreviation for the three cone fundamentals [$l(\lambda)$, $m(\lambda)$ and $s(\lambda)$] per individual, and can be written as a function with 10 parameters. The parameter a represents the age of the observer and v the visual angle (in degrees) of the stimulus. The age parameter was included in the CIEPO06 model to account for the difference in the ocular media between observers of different age groups, and does not necessarily correspond to the age of the real Stiles & Burch observers [117]. The other 8 parameters represent individual characteristics of the eye: d_{lens} is the deviation (in percentage) from the average lens pigment density; d_{macula} is the deviation (in percentage) in macular pigment density; $d_L, d_M,$ and d_S are deviations (in percentage) from the average peak optical density of the L -, M - and S -cone photopigments, respectively, and, s_L, s_M and s_S are shifts (in nm) from the average λ_{max} of L -, M -, and S - cone photopigments, respectively.

Our goal is to estimate the values of the parameters of Equation 7.7 for each participant from the isoluminance ratios measured in Experiment 1. These isoluminance ratios by definition mean that the sensation luminance is equal for the two extreme colors C_1 and C_2 of the corresponding chromatic modulation. Thus, for all the 60 isoluminance ratios, SL_1 equals SL_2 , where SL_i is obtained from Equation 7.2. In this equation, $v(\lambda)$ is usually approximated by a linear combination of the L -cone [$l(\lambda)$] and M -cone [$m(\lambda)$] spectral sensitivities: $v(\lambda) = l(\lambda) + w \times m(\lambda)$ [133], where w is the relative weight between L - and M -cone spectral sensitivities. Hence, Equation 7.2 can be rewritten as:

$$SL_i = K_m \int SPD_{C_i} (l(\lambda) + w \times m(\lambda)) d\lambda \{i = 1, 2\} \quad (7.8)$$

where SPD_{C_i} refers to the spectral power distribution of C_i . This spectral power distribution

can be obtained by transforming the u'_{10° , v'_{10° , Y -values (i.e., luminance) of C_i into rgb-values, using the calibration model of the lighting system. Then SPD_{C_i} is estimated by:

$$SPD_{C_i} = r_i \times R(\lambda) + g_i \times G(\lambda) + b_i \times B(\lambda) \{i = 1, 2\} \quad (7.9)$$

where $R(\lambda)$, $G(\lambda)$ and $B(\lambda)$ represent the spectra of the measured primaries of the system.

The cone fundamentals $l(\lambda)$ and $m(\lambda)$ in Equation 7.8 depend on the parameters of Equation 7.7. The value of w in Equation 7.8 is uncertain [130]. Sharpe et al. [123] proposed an average value of 1.55, but indicated that the value actually depends on the status of chromatic adaptation, on the viewing conditions and on the individual. For example, the value ranged between 0.56 to 17.75 among 40 observers [130]. Therefore, we treat w as an additional parameter in the model, instead of a constant value. Moreover, to keep the weight w comparable with literature, the cone spectral sensitivities [i.e., $l(\lambda)$, $m(\lambda)$ and $s(\lambda)$] were all normalized to unity peaks, as in [133], rather than being normalized to equal area as in [5].

To obtain the parameters of the model, we want to minimize the error function ϵ for each participant, being:

$$\epsilon = \sum_{i=1}^{n=60} \left| \log \left(\frac{SL_2}{SL_1} \right) \right| \quad (7.10)$$

For each of the fifteen participants, the visual angle was set to 10 degrees (i.e., the stimulus size in the experiment). Sarkar et al. [117] already pointed out that because of random observer variability, the best predicted model was not always obtained using the real age of the observer. In line with that observation, we treated the age parameter a as a free parameter. Since $L + M$ is widely accepted to represent the luminance channel, we omitted the S -cone contribution to the luminance and we set the two related variables (i.e., d_S and s_S) to zero. Thus, the ICO model is rewritten as follows:

$$lms_{ICO} = f(a, 10, d_{lens}, d_{macula}, d_L, d_M, 0, s_L, s_M, 0, w) \quad (7.11)$$

By combining Equations 7.8 to 7.11, the error function ϵ can be expressed in terms of the variables of Equation 7.11. We adopted the MATLAB optimization toolbox (i.e., *fmincon* solver) to estimate the seven parameters of Equation 7.11 for each of the fifteen participants. In the ICO model, each physiological parameter was assumed to follow a normal distribution. Thus, in the optimization procedure, each parameter was constrained to be within three standard deviations (± 3 SD) from the mean of the corresponding normal distribution. We repeated the procedure 1000 times, each time with a different random starting vector, in order to avoid the risk of finding a local minimum. The fit with the smallest value of ϵ was taken as the best function to describe the individual cone fundamentals.

Since taking into account individuals' deviations from the standard observer should improve the prediction of their perception, we expect to find a lower minimum for the error function ϵ when expressing isoluminance in terms of SL_i using the estimated individual cone fundamentals than when expressing isoluminance in terms of L_i using the CIE 10° standard cone fundamentals. Table 7.2 gives the error ϵ for the two different methods and the relative difference between these methods per participant. As can be seen, taking individual cone fundamentals into account results indeed in a better description of the individual isoluminance ratios. For some participants the difference is negligible, but for others the error is decreased by 50%.

Table 7.2: The error ϵ , as given in Equation 7.10, calculated once using the CIE 10° standard observer cone fundamentals (i.e., ϵ_{CIE}), and once using the best-fitted ICO-model parameters for the individual cone fundamentals (ϵ_{ICO}). The last column gives the difference between both methods in a percentage improvement, calculated as $100 \times (\epsilon_{CIE} - \epsilon_{ICO}) / \epsilon_{CIE}$.

Participant Number	ϵ CIE standard 10° observer	ϵ best-fit ICO	Difference (%)
1	0.0113	0.0111	-1.58
2	0.0293	0.0178	-39.39
3	0.0195	0.0136	-30.46
4	0.0248	0.016	-35.63
5	0.0216	0.019	-11.81
6	0.0238	0.0205	-13.85
7	0.0235	0.0223	-5.22
8	0.0121	0.0112	-7.48
9	0.0194	0.0178	-8.29
10	0.0175	0.0173	-1.06
11	0.0283	0.0205	-27.73
12	0.0162	0.0129	-20.42
13	0.0336	0.0162	-51.88
14	0.0175	0.0136	-22.3
15	0.0267	0.0133	-50.18

Figure 7.6(a) presents the estimated individual cone fundamentals of the fifteen participants, including the S-cone fundamentals. Although we didn't use these S-cone fundamentals in our fit to the isoluminance ratios (i.e., d_S and s_S were set at 0), still the S-cone fundamentals are slightly different because of age, d_{lens} and d_{macula} . Figure 7.6(b) shows the normalized individual luminous efficiency function $v(\lambda)$, calculated by $l(\lambda) + w \times m(\lambda)$. Since the procedure to estimate the individual cone fundamentals was based on the ratio in sensation luminance between two colors (C_1 and C_2) instead of on absolute values, $v(\lambda)$ is shown in normalized terms. The value of w ranged between 0.23 and 2.5, which is in line with the findings of Luria and Neri [82] and Sharpe et al. [123].

7.4. EXPERIMENT 2: SENSITIVITY TO CHROMATIC TEMPORAL MODULATIONS

7.4.1. PARTICIPANTS

Since the second experiment was intended to measure the visibility threshold for chromatic flicker at multiple frequencies for the above-mentioned color base points and modulations directions, it became very time consuming. Therefore, we limited ourselves to three of the fifteen participants of Experiment 1. These participants were male with an age of 23, 23 and 24 years. One of the three had corrected vision and was wearing untinted contact lenses during the experiment. The participants were selected based on their average isoluminance ratio: Participant 1 had a medium isoluminance ratio ($M = 1.0012$, $SD = 0.0215$), while Participant 2 had a relatively low average isoluminance ratio ($M = 0.9870$, $SD = 0.0154$) and Participant 3 a relatively high isoluminance ratio ($M = 1.0114$, $SD = 0.0353$). Since these

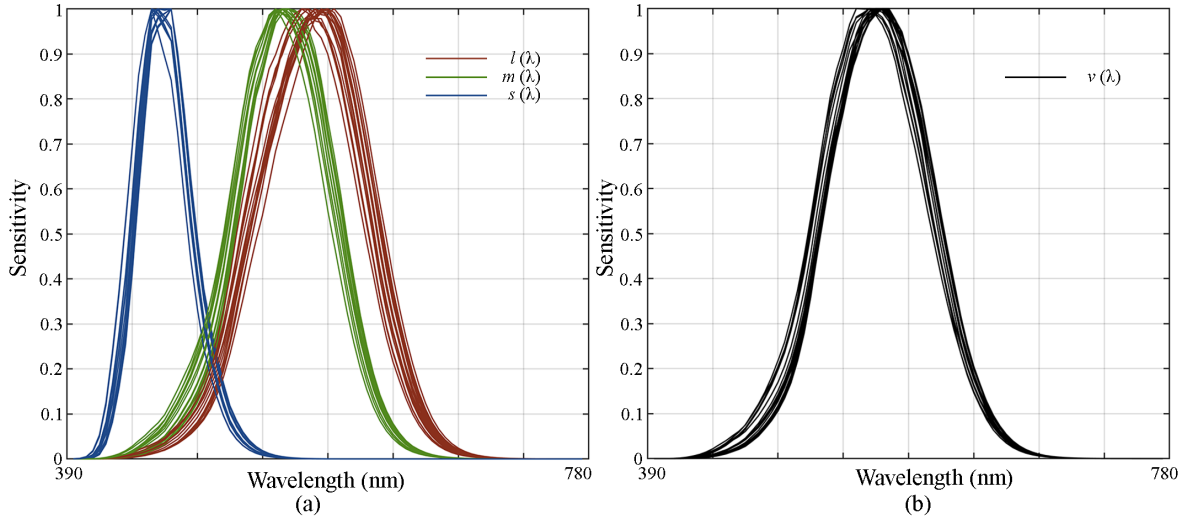


Figure 7.6: (a) The individual cone fundamentals [$s(\lambda)$, $m(\lambda)$, and $l(\lambda)$] of the fifteen observers based on the minimization of the error function given in Equation 7.9, and (b) the individual luminous efficiency function $v(\lambda)$ of the fifteen observers using their individual $l - m$ weight.

participants are a good representation of the variance in the isoluminance ratio measured in Experiment 1, their results are expected to give insight into the impact of individual differences in cone fundamentals on the sensitivity to chromatic modulations.

7

7.4.2. STIMULI

According to the results of **Chapter 6**, two frequencies should be sufficient to describe a TCSE. However, we decided to measure the contrast sensitivity for three frequencies (i.e., 2 Hz, 8 Hz and 20 Hz) in order to be able to check the exponential behavior again. The isoluminance ratios per participant obtained in Experiment 1 were used to ensure that each stimulus was isoluminant. However, these ratios were measured at just one amplitude of the modulation, while we needed them at more amplitudes to measure the visibility threshold for chromatic flicker. In a pilot study, where three amplitudes were used for the isoluminance experiment (i.e., $\Delta(u'_{10^\circ}, v'_{10^\circ}) = 0.01, 0.03$ and 0.05), a linear relation was found between the amplitude of the chromatic modulation and the logarithm of the corresponding isoluminance ratio. Therefore, the following equation was used to determine the individual luminance ratio per amplitude:

$$LR = LR_{iso}^{\frac{A}{A_{iso}}} \quad (7.12)$$

where A is the amplitude of the chromatic modulation that we need for Experiment 2, $A_{iso}=0.05$ (i.e., the amplitude that was used in Experiment 1), LR_{iso} is the isoluminance ratio measured in Experiment 1, and LR is the corresponding amplitude-corrected isoluminance ratio needed in this second experiment. In total, the visibility threshold for chromatic flicker was measured for 180 conditions (i.e., 15 Base Color \times 4 Modulation Direction \times 3 Frequency).

7.4.3. PROCEDURE

The task of the participant was to find the amplitude A at which flicker was just visible. They could increase the amplitude with large steps by pressing the Up-key and with small steps by

pressing the Right-key, and decrease the amplitude with large steps using the Down-key and with small steps using the Left-key. The amplitude was changed according to the following equation:

$$\frac{A_{n+1}}{A_n} = 1.05^{\pm\alpha} \quad (7.13)$$

where A_n refers to the current modulation amplitude, A_{n+1} to the modulation amplitude of the next stimulus, and α equaled 1 for the fine tuning (i.e., a change of about 5%) and 5 for the rough tuning (i.e., a change of about 27%). When the amplitude reached the system's lower boundary (i.e., $\Delta(u'_{10^\circ}, v'_{10^\circ}) = 0.00004$) or higher boundary (i.e., $\Delta(u'_{10^\circ}, v'_{10^\circ}) = 0.05$), a warning sound was played to indicate that the amplitude could not be further decreased or increased, respectively. When the participant reached the stimulus that appeared to have just visible flicker, s(he) was instructed to look at the final stimulus for one or two seconds before deciding to press the Enter-key to finish the trial. A beep sound was played to indicate the start of a new trial. To correct for a possible error of anticipation, all 180 conditions were presented twice: starting either at a modulation amplitude of $\Delta(u'_{10^\circ}, v'_{10^\circ}) = 0.0004$ (i.e., invisible flicker) or at a modulation amplitude of $\Delta(u'_{10^\circ}, v'_{10^\circ}) = 0.05$ (i.e., noticeable flicker). As such, the experiment consisted of 360 trials in total.

The stimuli were divided into fifteen different blocks with each block having a fixed base color, to keep the chromatic adaptation point as constant as possible. Within a block, all 24 trials (i.e., 4 *Modulation Direction* \times 3 *Frequency* \times 2 *Initial Amplitude*) with the same base color were randomized. Before the start of a new base color block, a two-minute period was included for the participants to adapt to the new base color at a luminance level of 37.5 cd/m². We presented at most two base color blocks in a session. Between two sessions, a minimum break of thirty minutes was mandatory to minimize fatigue and eye-strain of the participant. The sessions were distributed over different days. The fifteen base colors were assessed in a fixed order (i.e., $BC_1, BC_2, \dots, BC_{15}$). In total, the experiment lasted about four hours per participant.

7.4.4. RESULTS

DETECTION THRESHOLDS IN $u'_{10^\circ}u'_{10^\circ}$ CHROMATICITY DIAGRAM

The experiment resulted in 360 values (i.e., 15 *Base Color* \times 4 *Modulation Direction* \times 3 *Frequency* \times 2 *Initial Amplitude*) of the detection threshold expressed in $\Delta(u'_{10^\circ}u'_{10^\circ})$ per participant. First, the thresholds of the two initial amplitudes were averaged, and that result is plotted in the $u'_{10^\circ}u'_{10^\circ}$ chromaticity diagram in Figure 7.7, with a graph per frequency. Within such a graph, the thresholds per modulation direction with the same base color are connected to form a polygon per participant. Hence, Figure 7.7 shows three polygons around each base color, one per participant. We fitted ellipsoids through the eight points of each polygon, but this yielded inaccurate results for some conditions. To make the thresholds better visible, the polygons in Figure 7.7 have a zoom factor of 6, 5 and 1.2 at 2 Hz, 8 Hz and 20 Hz respectively.

Figure 7.7 shows that detection thresholds expressed in $u'_{10^\circ}u'_{10^\circ}$ chromaticity diagram depend on (1) *Base Color*, since the size of the polygons is not identical everywhere in the color space, (2) *Modulation Direction*, as the polygons are not octagon-shaped, (3) *Frequency*, since the size of the polygons differs between the three graphs, and (4) *Participant*, as the polygons seldomly overlap for the same frequency and base color. To confirm the above observations, an UNIANOVA analysis was performed with *Detection Threshold* (expressed in

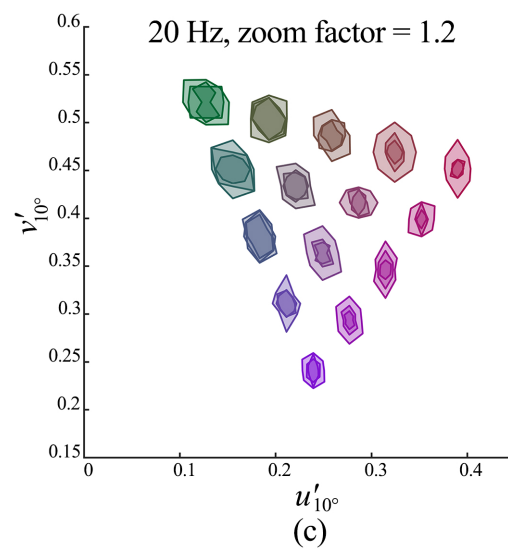
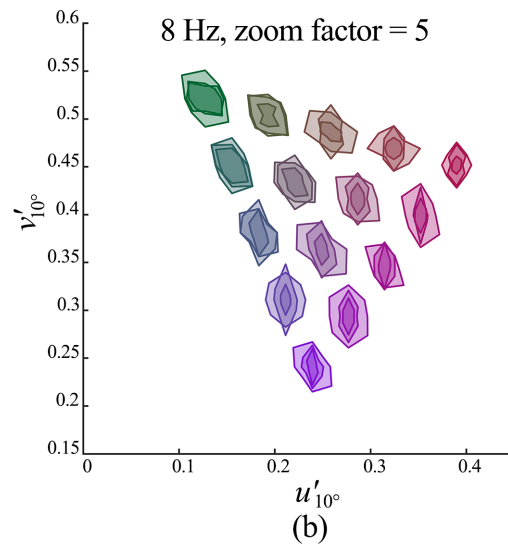
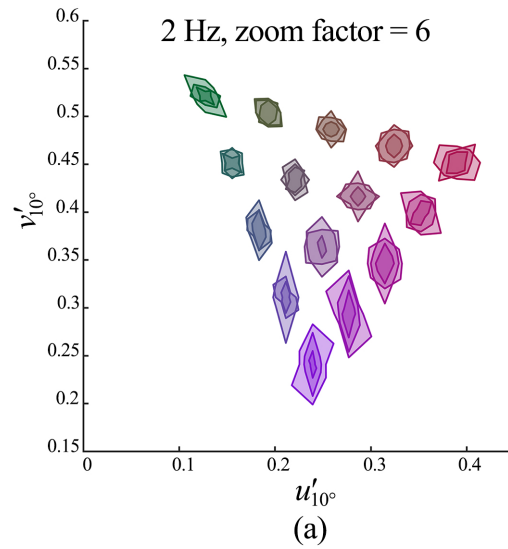


Figure 7.7: Polygons, indicating the visibility threshold for chromatic flicker per modulation direction, for the fifteen base colors and for the three participants at different temporal frequencies: (a) Frequency = 2 Hz; (b) Frequency = 8 Hz; (c) Frequency = 20 Hz. For the purpose of better visual presentation, the polygons have a zoom factor of 6, 5, 1.2 for the three frequencies respectively. The values of these zoom factors are manually tuned.

$u'_{10^\circ} u'_{10^\circ}$) as the dependent variable, *Base Color*, *Modulation Direction* and *Frequency* as the fixed independent variables and *Participant* as a random variable. The 2-way interaction effects of the fixed independent variables were included as well. The analysis indeed revealed significant effects of *Base Color* ($F(14, 442) = 6.1, p < 0.001$, partial $\eta^2 = 0.16$), *Modulation Direction* ($F(3, 442) = 40.4, p < 0.001$, partial $\eta^2 = 0.22$), *Frequency* ($F(2, 442) = 938.4, p < 0.001$, partial $\eta^2 = 0.81$) and *Participant* ($F(2, 442) = 61.6, p < 0.001$, partial $\eta^2 = 0.22$) on the detection thresholds. Besides, the interaction between *Base Color* and *Modulation Direction* was not significant ($F(42, 442) = 1.3, p = 0.1$, partial $\eta^2 = 0.11$). While the interactions between *Modulation Direction* and *Frequency* ($F(6, 442) = 8.6, p < 0.001$, partial $\eta^2 = 0.10$), *Base Color* and *Frequency* ($F(28, 442) = 6.2, p < 0.001$, partial $\eta^2 = 0.28$) were significant. These findings indicate that the $u'_{10^\circ} u'_{10^\circ}$ color space is not locally uniform, because of the effect of modulation direction, and also not globally uniform, because of the effect of base color. In addition, the shape of the polygons depends on frequency, indicating the need for a frequency-dependent model.

DETECTION THRESHOLDS IN LMS AND DKL

The $u'_{10^\circ} u'_{10^\circ}$ color space was chosen to define our stimuli. It was designed to be a more spatial uniform color space compared to CIE 1931 chromaticity diagram and temporal uniformity was not considered. We further analyzed the data in the physiologically based space (i.e., LMS cone excitation space). In **Chapter 3** and **Chapter 4**, it was shown that also DKL space [12] might be a good option for improving the temporal uniformity. Therefore, our data are analyzed in DKL as well.

In Section 7.3.4 we showed that the isoluminance ratios of the individual participants can be described better by using the (estimated) cone fundamentals of the individual instead of those of the standard observer. We expect that the individual cone fundamentals can also be used to minimize or even eliminate the variation in the sensitivity to chromatic temporal modulations between individuals. Therefore, we transformed the $\Delta(u'_{10^\circ}, v'_{10^\circ})$ values of the threshold into a contrast measure in LMS and DKL, once based on the standard observer using Stockman-Sharpe (SS) cone fundamentals and once based on the cone fundamentals optimized for individual colorimetric observers (ICO).

Given a detection threshold expressed in $\Delta(u'_{10^\circ}, v'_{10^\circ})$, the two extreme colors C_1 and C_2 can be determined using Equation 7.5. Next, the spectral power distribution of C_1 and C_2 can be obtained by transforming the $u'_{10^\circ}, v'_{10^\circ}, Y$ -values (i.e., luminance) into rgb -values, using the calibration model of the lighting system, and then using Equation 7.9. Then, $LMS_{C_i ICO}$ ($i = 1, 2$) and $LMS_{C_i SS}$ ($i = 1, 2$) can be calculated as follows:

$$\begin{cases} LMS_{C_i ICO} = \int SPD_{C_i(\lambda)} \times lms_{ICO} d\lambda \\ LMS_{C_i SS} = \int SPD_{C_i(\lambda)} \times lms_{SS} d\lambda \end{cases} \quad (7.14)$$

where lms_{ICO} and lms_{SS} are the cone fundamentals (i.e., with a dimension of n -by-3, where n is 79—from 390 nm to 780 nm with a 5-nm resolution) and SPD_{C_i} ($i = 1, 2$) represent the spectral power distribution (i.e., with a dimension of 1-by- n , where n is 79) of the two extreme colors.

$$\left\{ \begin{array}{l}
C_{SS-LMS} = \sqrt{\frac{\Delta L_{SS}^2}{L_{SS-BC}} + \frac{\Delta M_{SS}^2}{M_{SS-BC}} + \frac{\Delta S_{SS}^2}{S_{SS-BC}}} \\
C_{ICO-LMS} = \sqrt{\frac{\Delta L_{ICO}^2}{L_{ICO-BC}} + \frac{\Delta M_{ICO}^2}{M_{ICO-BC}} + \frac{\Delta S_{ICO}^2}{S_{ICO-BC}}} \\
C_{SS-DKL} = \sqrt{\left[\Delta(L_{SS} - M_{SS})\right]^2 + \left[\Delta[S_{SS} - (L_{SS} + M_{SS})]\right]^2} \\
C_{ICO-DKL} = \sqrt{\left[\Delta(L_{ICO} - M_{ICO})\right]^2 + \left[\Delta[S_{ICO} - (L_{ICO} + w \times M_{ICO})]\right]^2}
\end{array} \right. \quad (7.15)$$

where Δ refers to the difference in the corresponding cone spectral sensitivities between the two extreme colors of the modulation, and w is the relative weight between L - and M -cone spectral sensitivities, and its value is identical as in Equation 7.8 and Equation 7.11. L_{SS-BC} , M_{SS-BC} , and S_{SS-BC} represent the L -, M -, and S -cone excitation induced by the base color calculated using the SS cone fundamentals, while L_{ICO-BC} , M_{ICO-BC} , and S_{ICO-BC} represent the L -, M -, and S -cone excitation induced by the base color calculated using the ICO cone fundamentals.

CIRCULARITY – LOCAL UNIFORMITY

As described above, STRESS-based metrics were proposed [112] to quantify (a) the local uniformity of the color space by taking the ratios of the major and minor semi-axis of the color discrimination ellipses as input, and (b) the global uniformity by taking the areas of the ellipses as input. Those two metrics rely on fitting ellipses, thus, the accuracy is largely dependent on the goodness-of-fit of the ellipse-fitting algorithm. In our study, we fitted ellipsoids through the eight points of each polygon, but this yielded inaccurate results for some conditions. So, we propose a more general way of computing circularity and homogeneity by eliminating the limitation of fitting ellipses.

Circularity is the degree in which the detection threshold depends on the direction of the modulation in the color space. The deviation from circularity ($dCIR$) for a give participant (pp) and temporal frequency (f) is calculated as:

$$dCIR_{pp,f} = \frac{\sum_{b=1}^{n_b} \sum_{d=1}^{n_d} \left| \frac{C_{b,d} - \overline{(C_b)_d}}{\overline{(C_b)_d}} \right|}{n_d \times n_b} \quad (7.16)$$

where $C_{b,d}$ is the contrast measure of a certain modulation direction d around base color b ; $\overline{(C_b)_d}$ is the median of the contrast measures of the four modulation directions around a certain base color b ; n_b and n_d represent the number of base colors (i.e., 15) and the number of modulation directions (i.e., 4), respectively. The value of $dCIR$ ranges between zero and infinity. When $dCIR$ equals zero, the detection thresholds are independent of modulation direction at all base color points, which means that the color space is perfectly locally uniform. A higher $dCIR$ value indicates a lower degree of circularity. We calculated $dCIR$ for each of the four contrast measures defined in Equation 7.15 and for the detection thresholds calculated as $\Delta(u'_{10^\circ}, v'_{10^\circ})$. The results are shown in Figure 7.8.

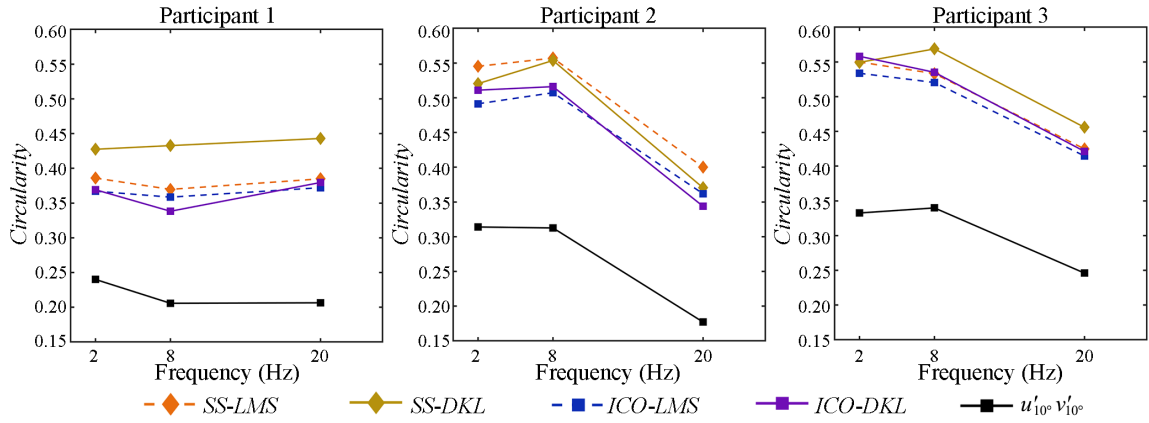


Figure 7.8: *Circularity* of the four contrast measures and of $\Delta(u'_{10^\circ}, v'_{10^\circ})$ as a function of temporal frequency for each participant.

Figure 7.8 shows that the circularity changes with frequency; for two of the three participants (i.e., Participant 2 and 3) the circularity peaks at the highest frequency (i.e., $dCIR$ value is the lowest). When comparing the two color spaces, it's hard to draw a consistent conclusion about whether the circularity is higher for the DKL-based color space or for the LMS-based color space, because it depends heavily on participant. In general, the use of individual cone fundamentals instead of standard cone fundamentals slightly improves the *circularity*. In contrast to our expectations, the physiologically based color spaces have worse local uniformity than the $u'_{10^\circ} v'_{10^\circ}$ color space. Thus, from this perspective, it seems to make no sense to look into cone fundamentals to describe a temporally uniform color space. However, we should not forget that the sinusoidal modulations used in the subjective experiment were designed in the $u'_{10^\circ} v'_{10^\circ}$ color space, which might explain the better results.

HOMOGENEITY – GLOBAL UNIFORMITY

Homogeneity is the degree in which the detection threshold depends on the location in the color space. The deviation from *homogeneity* (dH) for a give participant (pp) and temporal frequency (f) was calculated as:

$$dH_{pp,f} = \frac{\sum_{d=1}^{n_d} \sum_{b=1}^{n_b} \left| \frac{C_{d,b} - \overline{(C_d)_b}}{\overline{(C_d)_b}} \right|}{n_b \times n_d} \quad (7.17)$$

where $C_{d,b}$ is the contrast measure of a certain modulation direction d around base color b ; $\overline{(C_d)_b}$ is the median of the contrast measures at fifteen base colors along a certain modulation direction d ; n_b and n_d represents the number of base colors (i.e., 15) and the number of modulation directions (i.e., 4), respectively. Similar to deviation from *circularity* ($dCIR$), the value of dH ranges between zero and infinity. When dH equals zero, the detection thresholds are independent of base colors along all modulation directions, which means that the color space is perfectly globally uniform. A higher dH value indicates a lower degree of homogeneity. Also here we calculated the homogeneity dH for all four contrast measures defined in Equation 7.15 and for the detection thresholds calculated as $\Delta(u'_{10^\circ}, v'_{10^\circ})$, the results of which are shown in Figure 7.9.

As can be seen in Figure 7.9, frequency has a different effect on homogeneity for the three participants, and the homogeneity shows a different trend for the LMS-based color space

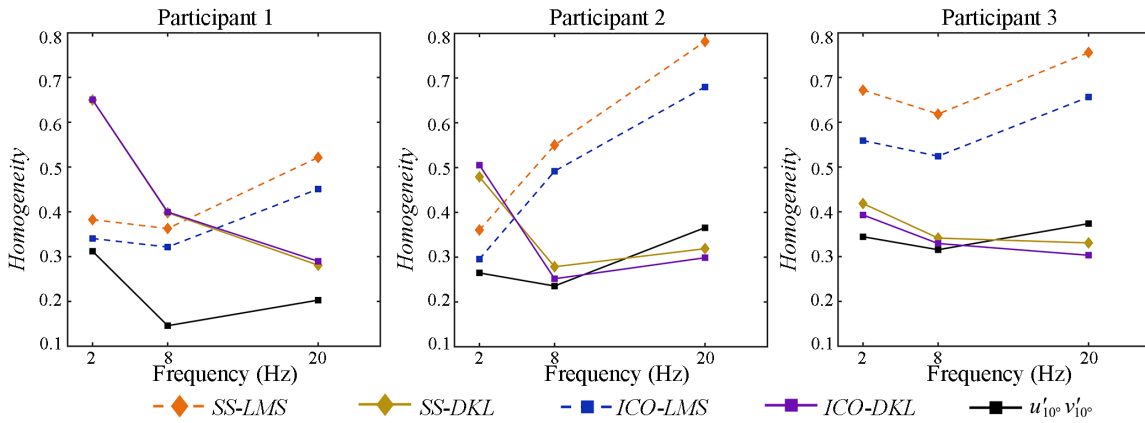


Figure 7.9: The *homogeneity* of the four contrast measures and of $\Delta(u'_{10^\circ}, v'_{10^\circ})$ as a function of temporal frequency for each participant.

than for the DKL-based color space. For participant 1, homogeneity improves with frequency in the DKL-based color space; homogeneity is the lowest at the highest frequency and higher at the low and intermediate frequency in the LMS-based color space. For participant 2 homogeneity decreases with frequency in the LMS-based color space, and is slightly better at the intermediate frequency than at the highest frequency in the DKL-based color space. For participant 3 homogeneity is consistently better in the DKL-based color space than in the LMS-based color space. In general, the DKL-based color space shows better homogeneity than the LMS-based color space at higher temporal frequencies. The influence of using individual cone fundamentals is smaller in the DKL-based color space than in the LMS-based color space. Although the *homogeneity* at the highest temporal frequency is highest when using ICO cone fundamentals in DKL space for both Participants 2 and 3, in general similar or better results in terms of *homogeneity* are obtained in the $u'_{10^\circ} v'_{10^\circ}$ color space. So, again in contrast to our expectations, also for the global uniformity of the temporal color space the use of cone fundamentals seems to have no added value with respect to the $u'_{10^\circ} v'_{10^\circ}$ color space.

When combining Figure 7.8 and Figure 7.9, it seems that using a DKL-based contrast measure gives the best results in terms of both *circularity* and *homogeneity* among the four contrast measures in Equation 7.15. However, this measure in general yields worse results than simply using the $u'_{10^\circ} v'_{10^\circ}$ color space. Still, none of these measures show a consistent trend of the flicker detection threshold with frequency for all three participants.

FREQUENCY-CONSISTENCY

As shown above, the thresholds strongly depend on frequency. In **Chapter 6**, we showed that the relation between frequency and contrast sensitivity (S) could be described by a single exponential function for all tested combinations of base color and modulation direction. Such a single exponential function can be written as:

$$\ln S = \beta_1 f + \beta_0 \quad (7.18)$$

with β_1 being the slope and β_0 the intercept of the function. The question is how these parameters change with base color and modulation direction and how much they vary between individuals. Therefore, we fitted Equation 7.18 to the data of each participant for

each combination of base color and modulation direction separately, for all the four contrast measures computed by Equation 7.15. The sensitivity S is the reciprocal of the contrast measure, as shown in Equation 7.3. The average R^2 -values using four contrast measures were all higher than 0.91. Here we take SS-LMS contrast as an example. The values of the slopes and intercepts for all conditions and all three participants are presented in Figure 7.10. This figure can be compared to Figure 6.5 in Chapter 6, where the authors give a similar graph, but then for sensitivities calculated in the $u'_{10^\circ}v'_{10^\circ}$ color space.

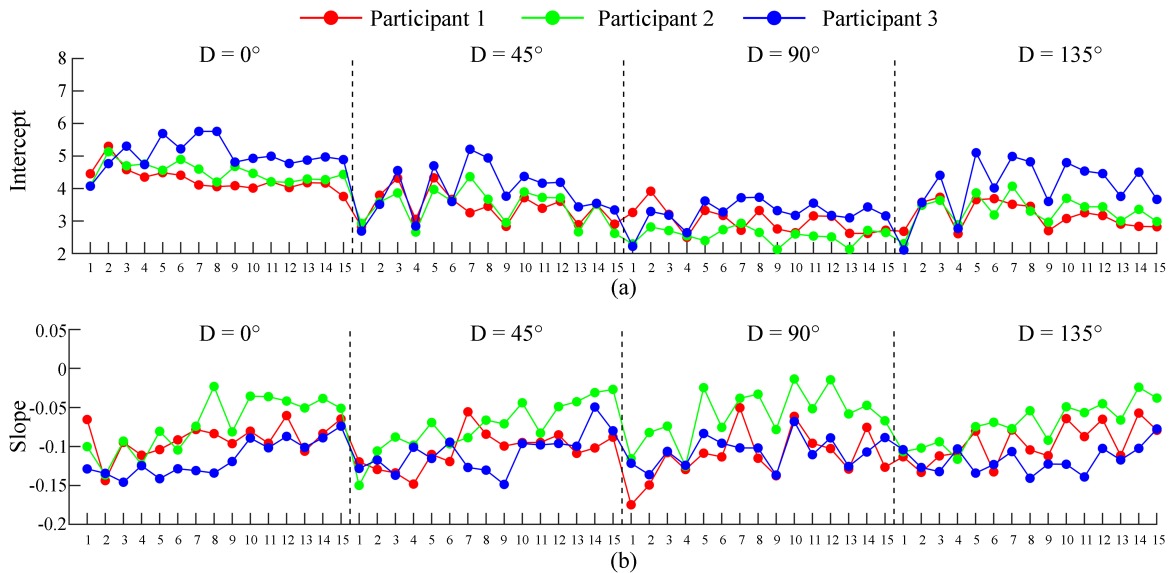


Figure 7.10: (a) Intercepts of the TCSFs of the three participants for all conditions, and, (b) Slopes of the TCSFs of the three participants for all conditions. In both graphs, the intercepts and slopes were computed using SS-LMS.

The R^2 -values of the fits (i.e., a linear fit with three frequencies) were on average 0.94 (SD = 0.11) for Participant 1, 0.86 (SD = 0.21) for Participant 2, and 0.93 (SD = 0.09) for Participant 3. This again shows that the effect of frequency on temporal contrast sensitivity can be modeled by a single function with variable parameters. From all the 180 fits (i.e., 3 Participant \times 15 Base Color \times 4 Modulation Direction), only 11 fits had an R^2 -value smaller than 0.6, which usually indicates a poor fit. Eight of these 11 poor fits could be attributed to base colors BC_{10} to BC_{15} measured with Participant 2. For those fits with a low R^2 , the measured contrast sensitivity at the middle frequency (i.e., 8 Hz) was higher than the contrast sensitivity measured at the low frequency (i.e., 2 Hz). Since only a low percentage of fits (i.e., 6%) resulted in a low R^2 -value, we did not exclude these fits from further analyses.

Figure 7.10 illustrates that some general trends may be observed in the values of the slope and intercept for the three participants, but there are also obvious individual differences. Since the focus of this chapter is not on quantifying how the slope and intercept vary over the color space, we do not report statistical analyses, like in Chapter 6. In the next section we look further into the differences between participants.

INDIVIDUAL DIFFERENCES

Clearly people differ in the effect of base color, modulation direction and frequency on the temporal contrast thresholds. However, keep in mind that the degree of individual differences does not say anything about the (local and global) uniformity of the color space.

On the other hand, we do expect individual differences to be smaller when we take into account individually adapted cone spectral sensitivities. In addition, the variability over participants may be different for different color spaces, hence in our analysis we include both the LMS-based and DKL-based color spaces.

The overall difference between individuals can be quantified as follows:

$$\Delta_{ID} = \frac{\sum_{f=1}^{n_f} \sum_{d=1}^{n_d} \sum_{b=1}^{n_b} \sum_{p=1}^{n_p} \left| \frac{C_{f,d,b,p} - (\widetilde{C_{f,d,b}})_p}{(\widetilde{C_{f,d,b}})_p} \right|}{n_p \times n_b \times n_d \times n_f} \quad (7.19)$$

where $C_{f,d,b,p}$ is the contrast measure for a certain modulation direction d around base color b of participant p at temporal frequency f ; $(\widetilde{C_{f,d,b}})_p$ is the median of the contrast measures of the three participants for a certain modulation direction d around base color b at temporal frequency f ; n_p , n_b , n_d and n_f represent the number of participants (i.e., 3), the number of base colors (i.e., 15), the number of modulation directions (i.e., 4), and the number of temporal frequencies (i.e., 3), respectively. The value of Δ_{ID} ranges between zero and infinity and the value of zero means that all three participants are identical. A larger value of Δ_{ID} indicates larger differences between these individuals.

Since there were only three participants in Experiment 2, $(\widetilde{C_{f,d,b}})_p$ is identical to one of the participants' contrast measure at the corresponding base color, modulation direction and temporal frequency. To further investigate the individual differences between any two participants—where the median is the mean of two data points—Equation 7.19 can be adapted as follows:

$$\Delta_{ID(p_i,p_j)} = \frac{\sum_{f=1}^{n_f} \sum_{d=1}^{n_d} \sum_{b=1}^{n_b} \left| \frac{C_{f,d,b,p_i} - C_{f,d,b,p_j}}{C_{f,d,b,p_i} + C_{f,d,b,p_j}} \right|}{n_b \times n_d \times n_f} \quad (7.20)$$

where $\Delta_{ID(p_i,p_j)}$ represents the difference between participant p_i and p_j ($i=1,2,3; j=1,2,3$); C_{f,d,b,p_i} and C_{f,d,b,p_j} represent the contrast measure of participant p_i and p_j , respectively, at temporal frequency f for modulation direction d around base color b ; n_b , n_d , and n_f represent the number of base colors (i.e., 15), the number of modulation directions (i.e., 4), and the number of temporal frequencies (i.e., 3), respectively.

Figure 7.11 shows the calculation of the sum of the individual differences as given by Equation 7.19, as well as the differences for each combination of two participants, as calculated with Equation 7.20. It illustrates that participant 1 and participant 2 are most different, while participant 2 and participant 3 are most similar.

Overall individual differences are smallest when using $C_{ICO-LMS}$ as the contrast measure, although these differences are only slightly higher when using $\Delta(u'_{10^\circ BC}, v'_{10^\circ BC})$. For both color spaces, using individual cone fundamentals resulted in smaller individual differences than using standard observer cone fundamentals, as we would expect.

7.5. DISCUSSION

In this study, we performed two psychophysical experiments to measure the detection thresholds of isoluminant chromatic flicker. In Experiment 1, individual isoluminance ratios were measured for fifteen participants separately, using a technique similar to HFP. In Experiment 2, the detection thresholds of isoluminant chromatic flicker were measured at fifteen base colors, four modulation directions and three temporal frequencies. The isoluminance

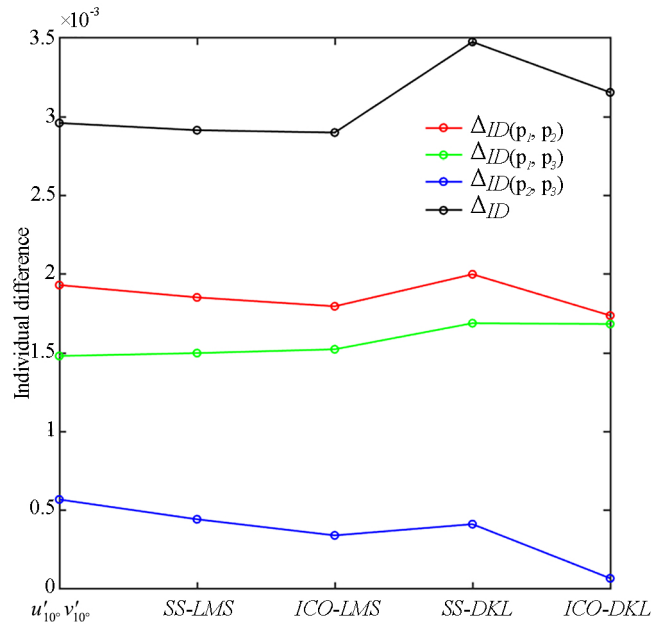


Figure 7.11: Overview of individual differences calculated with different contrast measures for all participants together (black curve) as well as for each combination of two participants (red, green and blue curve).

7

ratios were used to model individual cone fundamentals based on the ICO model, and we showed that these individual cone fundamentals better predict the isoluminance ratios than standard observer cone fundamentals. The chromatic flicker detection thresholds were used to explore three aspects: 1) *circularity*, which describes the local uniformity of a color space; 2) *homogeneity*, which describes the global uniformity of a color space; and 3) *frequency-consistency*, which describes the relation between detection thresholds and temporal frequency. Furthermore, circularity and homogeneity were systematically evaluated in commonly used cone-excitation color spaces, such as LMS and DKL. Besides the specific choice of contrast measures used in this chapter (i.e., as listed in Equation 7.15), further variations can be thought of and are worth exploring in future research.

7.5.1. ESTIMATING INDIVIDUAL CONE FUNDAMENTALS

The isoluminance experiment was performed for two reasons. First, to be able to measure TCSFs for isoluminant chromatic flicker stimuli. Second, to estimate individual cone fundamentals. Luminance is defined for an average observer, thus it is different for each real observer. At medium-high temporal frequencies, human observers are more sensitive to luminance flicker than chromatic flicker [16]. Thus, for measuring the detection thresholds of chromatic flicker at higher temporal frequencies, luminance flicker should be eliminated, which implies that the stimuli have to be configured for each participant experimentally. Using HFP to configure isoluminant stimuli has been a common practice to achieve this goal.

HPF is considered as a reliable and accurate method to measure a subject's luminous efficiency function [72]. Thus, we employed this method to configure individual isoluminance in this study. A modulation frequency of 25 Hz was used, as it is a commonly used frequency for HFP (e.g. [123]), at which human observers are quite sensitive to luminance flicker. A

fixed amplitude A of $\Delta(u'_{10^\circ}, v'_{10^\circ}) = 0.05$ was adopted because it often gives a very visible chromatic flicker, and at the same time all stimuli remained within the lighting system's gamut. In a pilot study, three amplitudes were used for the isoluminance experiment (i.e., $\Delta(u'_{10^\circ}, v'_{10^\circ}) = 0.01, 0.03$ and 0.05), and a linear relation was found between the isoluminance ratio and the amplitude of the chromatic modulation. Thus, the isoluminance ratios can be calculated with Equation 7.12 to generate isoluminant chromatic flicker stimuli for other amplitudes. Although this method minimizes luminance flicker, we cannot exclude that some residual luminance flicker may still have been perceived by the participants.

To estimate the individual cone fundamentals, we used the ICO model, which was extended based on CIEPO06. The ICO model allowed more freedom in the variations of cone fundamentals. We added an additional parameter w , which is the relative weight between L - and M -cone spectral sensitivities when adopting the ICO model. This weight may vary largely; for example, Stockman and Brainard [130] reported values between 0.56 and 17.75. Among our fifteen participants the parameter w varied between 0.23 and 2.5. In the optimization procedure for deriving individual cone fundamentals using the isoluminance ratios, we assumed that the S -cone did not contribute to the luminance channel, which is a widely accepted assumption. As a result, individual sensitivities of S -cones could not be derived. However, we cannot exclude that also the S -cones contributed to the luminance channel. For example, it is known that S -cones make a small, but robust contribution to luminance under conditions of intense long-wavelength adaptation [109]. Unfortunately, since adaptation could not always be fully assured in our experiment, we cannot estimate whether we had any S -cone contribution to luminance.

Thus, there are some limitations in using isoluminance ratios collected with HFP to estimate individual cone fundamentals. Other efforts to estimate individual cone fundamentals exist; for example, Andersen, Finlayson, and Connah [4] proposed to estimate individual cone fundamentals from their color matching functions. We did, however, not choose for this approach, since measuring individual color matching functions is considerably more tedious and time-consuming than using HFP. In a recent paper Lee et al. [74], the authors adopted luminance-based data to estimate individual color matches. In contrast to these approaches mentioned in the literature, our approach is more practical, and still results in a better-than-average estimation of the individual cone fundamentals. In addition, these individual cone fundamentals proved to be good at describing our luminance data. As our results showed, the ICO-based model was a better description of individual isoluminance ratios than the standard cone fundamentals. For a few participants the difference was negligible, but for others the error decreased by 50%.

7.5.2. QUANTITATIVE EVALUATION OF TEMPORAL COLOR SPACES IN TERMS OF LOCAL AND GLOBAL UNIFORMITY

Three participants who had a low, medium, and high average isoluminance ratio among the fifteen participants respectively continued with the experiment of collecting detection thresholds of chromatic flicker at fifteen base colors, four modulation directions and three temporal frequencies. The detection thresholds were expressed in different color spaces, i.e., $u'_{10^\circ} v'_{10^\circ}$ chromaticity diagram, the LMS-based color space, and the DKL-based color space. Furthermore, since individual cone fundamentals were estimated, the uniformity of the LMS and DKL color spaces was compared using Stockman-Sharpe (SS) cone fundamentals versus ICO-based fundamentals.

Circularity and *homogeneity* were the measures used to describe the local and global uniformity of a color space respectively. To define these measures, the first step is to select an appropriate contrast measure for the detection threshold. We evaluated four different measures, namely two in the LMS-based color space and two in the DKL-based color space. Since the stimuli were defined in the $u'_{10^\circ} v'_{10^\circ}$ chromaticity diagram, $\Delta(u'_{10^\circ}, v'_{10^\circ})$ was included as a contrast measure as well. The results demonstrated that both *circularity* and *homogeneity* vary with temporal frequency, and the effect of frequency is different for different observers. In addition, both measures showed that in general the DKL space is somewhat more uniform than the LMS space. This finding is consistent with **Chapter 6**, in which the DKL space was considered a good option to improve the temporal uniformity of a color space. When it comes to the comparison between ICO-based cone fundamentals and SS cone fundamentals, the ICO-based cone fundamentals improved the temporal uniformity of the color space both locally and globally, as expected.

However contrary to our expectations, the $u'_{10^\circ} v'_{10^\circ}$ color space has overall the best circularity and homogeneity. The latter may have three possible explanations. First, the chromatic stimuli used in the experiment were defined in the $u'_{10^\circ} v'_{10^\circ}$ color space, which means that the sinusoidal modulations in color followed a predictable path in this color space, but not necessarily in the LMS-based or DKL-based color spaces defined with standard or individual cone fundamentals. Ideally, if we wanted to analyze the chromatic flicker detection thresholds in a cone fundamental color space, we should have modulated the stimuli in that same color space. However, this is essentially a circular reasoning because we could not define isoluminant chromatic fluctuations in these cone fundamental color spaces, since the first experiment determined isoluminance ratios in the $u'_{10^\circ} v'_{10^\circ}$ color space. Second, the color spaces based on cone fundamentals should take into account the degree of chromatic adaptation of the human visual system when describing the perception of stimuli. We did not include chromatic adaptation in our descriptions, since it is not clear yet how to do that for chromatically modulated stimuli. Neglecting chromatic adaptation may thus be a reason why the description of the flicker detection thresholds does not result in a temporally uniform color space. Third, we used an implementation of the DKL-based color space, in which we added one weighting factor to weight the contribution of the M-cones with respect to the L-cones (see Equation 7.15). Further optimization of *circularity* and *homogeneity* may be achieved by adding other weighting factors, for example to weight the contribution ($L - M$) with respect to ($L + w \times M$). Whether this refinement is needed, will be investigated in our future research.

In summary, the quantitative measures of circularity and homogeneity are helpful for developing a (temporally) uniform color space in future, as they can be adopted to construct cost functions in maximizing uniformity measures. In addition, those measures can facilitate the comparison of uniformity across different color spaces.

7.5.3. FREQUENCY-CONSISTENCY

Frequency-consistency refers to the property that the effect of temporal frequency should be modeled consistently over all color points and modulation directions, despite that the actual value of the thresholds themselves may depend on temporal frequency. We found that the effect of frequency could be modeled by an exponential function (i.e., Equation 7.18) when taking the reciprocal of the contrast measures as representation of the contrast sensitivity. However, the intercepts and slopes were dependent on the base color and modulation

direction for all the three color spaces (i.e., the $u'_{10^\circ}v'_{10^\circ}$, LMS-, and DKL-based color spaces) tested above. This means that a more elaborate color space is needed to make its shape independent of modulation frequency.

7.5.4. INDIVIDUAL DIFFERENCES

It is known that there are many sources that lead to individual differences in human color vision, such as age [5] and the relative amount of different cones [38]. The three participants had a good representation of the variance in the isoluminance ratios measured in Experiment 1, thus their results are expected to give insight into the impact of individual differences in cone fundamentals on the sensitivity to chromatic modulations. Similar to circularity and homogeneity, a quantitative measure, defined in Equation 7.19, was proposed to estimate individual differences. A direct application of such a measure is to quantify the overall difference in sensitivity to chromatic temporal modulations between individuals. As expected, using individual cone fundamentals resulted in smaller individual differences than using a standard observer description.

7.5.5. ADAPTATION

Adaptation in the visual system is fundamental and ubiquitous [155]. Results of light/dark adaptation [70] and chromatic adaptation [108] have been widely documented. Adaptation in real-life vision is seldom complete, and the degree of adaptation depends on many things, including amongst others the field of view [83]. Currently, there is no good model to compute the degree of chromatic adaptation for temporally modulated colored stimuli. Thus, in our experiments, the stimuli were divided into different blocks with a fixed base color, to keep the chromatic adaptation point as constant as possible. Since the cone responses are related to the degree of adaptation, to better capture this nature, a contrast term that takes adaptation into consideration should be evaluated in future work.

The visual system is also known to adapt to the dynamics of light, such as flicker. Flicker adaptation has been described in literature [122], but the model for predicting it is still lacking. In our study, we presented an unmodulated two-minute stimulus of a fixed color before the measurement of the detection threshold of chromatic flicker in order to avoid that participants gradually adapted to flicker during the experiment. Nonetheless, uncontrolled effects of adaptation may have affected our experimental results, and attempts to disentangle better adaptation from the measurement of the perceived dynamics of colored light transitions may benefit the development of a temporally uniform color space.

7.6. CONCLUSION

In this study, individual cone fundamentals were estimated, and quantitative measures were proposed to describe the local and global uniformity (i.e., *circularity* and *homogeneity*) of color spaces and the impact of individual differences. The individual cone fundamentals were calculated based on the individual isoluminance ratios and were found to outperform the standard observer cone fundamentals in terms of describing the isoluminance data.

In general, the DKL color space is somewhat more uniform than the LMS color space and the ICO-based cone fundamentals improve both the *circularity* and *homogeneity*. However, contrary to our expectations, the $u'_{10^\circ}v'_{10^\circ}$ color space has overall the best *circularity* and *homogeneity*. In all color spaces, substantial individual differences were found: these

individual differences were largest in the $u'_{10^\circ}v'_{10^\circ}$ color space and smallest in the DKL color space. For both the LMS and DKL color space, individual differences were smaller when using ICO-based cone fundamentals than when using the standard observer cone fundamentals.

The data of the visibility thresholds for chromatic flicker confirmed the usefulness of an exponential function to model TCSFs, when expressing the sensitivity in the $u'_{10^\circ}v'_{10^\circ}$, LMS, and DKL color spaces. However, these color spaces are not able to eliminate the effect of base color and modulation direction on the intercepts and slopes of the TCSFs. Therefore, further modeling is needed to design a (more) perceptually uniform temporal color space that enables the design of dynamic color transitions that are pleasant for all observers.



General Discussion

LED lighting has been a game-changer in both academia and industry. Besides its widely known advantages, such as energy-saving and a small form factor, the fast temporal response of LEDs has brought both challenges and opportunities. LEDs enable researchers to generate well-controlled spatio-temporal stimuli for both fundamental and applied vision research, while companies can implement dynamic features in their products. However, a model that predicts how the human visual system responds to dynamic colored stimuli is still lacking. A new temporally uniform color space is of vital importance to ensure high-quality dynamic lighting, which means that the light should fulfill at least three criteria: (1) it should have an attractive color transition from one color to another, (2) the transition should appear to be smooth, and (3) it should have a controllable perceived speed.

In **Chapter 1** and **Chapter 2** we explained the challenges to build a temporal model and relevant background information is given. Establishing a temporally uniform color space is a task too big for a single PhD-project. However, with this long-term goal in mind, several psychophysical experiments were carried out, addressing specifically the last two aspects of high-quality dynamics - the perceived smoothness and speed of a colored light transition. As CIELAB is considered as one of the most (spatially) uniform color spaces available, it was chosen to generate most stimuli for our psychophysical experiments. Specifically, Delta-E-per-second in CIELAB was used to describe the modulation speed of a temporal color transition over time. In **Chapter 3** and **Chapter 4**, we investigated the effect of modulation direction (i.e., chroma and hue) and location in the color space (i.e. base color) on the perceived speed of a colored light transition defined in CIELAB under two different viewing conditions. Two psychophysical experiments were conducted that differed in spatial properties (such as the field-of-view and the surround of the stimulus), but both compared the perceived speed of a pair of temporal color transitions that changed in chroma or hue direction around five base colors. For each pair, subjects were asked to identify which transition appeared faster. The results were used to compute points of subjective equality (PSE) between the transitions. Despite differences between the two experiments, consistent results were obtained. We adopted the standard deviations of PSEs to measure the uniformity of different existing color spaces. It was no surprise that we did not find an existing color space able to predict the results. However, with simple modifications, both CIELAB-based

and cone-excitation based (such as DKL [12]) color spaces showed improved uniformity of the perceived speed.

The results in **Chapter 3** and **Chapter 4** confirmed that CIELAB is not uniform in predicting perceived speed of relatively slow temporal color transitions. In the rest of the chapters the same effects of base color and modulation direction were measured for relatively fast changes in color, perceived as chromatic flicker. To fasten data collection for the large amount of data points needed to build a uniform color space, five psychophysical methods were compared in terms of accuracy, precision, and efficiency in **Chapter 5**. In this comparison, we combined classical psychophysical methods with commonly used yes-no tasks and two-alternative forced-choice (2AFC) tasks, resulting in the following five methods: **1**) the classical 1-up-1-down staircase method with a yes-no task, **2**) the weighted 3-up-1-down staircase method with a 2AFC task, **3**) the method of constant stimuli with a 2AFC task, **4**) the method of adjustment without a reference, and **5**) the method of adjustment with a reference. The conclusion was that the method of adjustment without a reference had a good balance between accuracy and efficiency. In addition, this experiment gave first insights in the effect of base color and frequency on the detection threshold of chromatic flicker.

With the method selected from **Chapter 5**, we conducted the experiment presented in **Chapter 6** to extend the data on the effect of base color, modulation direction, and temporal frequency on the visibility threshold of chromatic temporal modulations. It was found that for any combination of base color and modulation direction, a single exponential function was sufficient to model the effect of temporal frequency on the reciprocal of the chromatic flicker visibility threshold (usually indicated as sensitivity). The resulted functions are commonly referred to as Temporal Contrast Sensitivity Functions (TCSFs). The intercepts and exponents of these functions depended on the specific base color and modulation direction, which further confirmed the non-uniformity of the color space in which sensitivity was defined. Moreover, for each condition, substantial individual differences were found not only in terms of absolute detection threshold, but also in terms of its change with temporal frequency (i.e., in terms of the intercept and exponent of the TCSF).

Although large individual differences are common in vision related psychophysical experiments, the development of a uniform color space aims for applications assuming an averaged (or standard) observer. Therefore, we recruited more participants for the experiment discussed in **Chapter 7**. We measured their sensitivity to chromatic temporal modulations, giving us insight in the variability in sensitivity. To disentangle individual differences in luminance flicker from chromatic flicker, we split the experiment in two parts. In the first part, we measured the isoluminance ratios (i.e., the luminance ratio between two colors, needed to make them equal in perceived luminance) of chromatic flicker stimuli using a technique similar to heterochromatic flicker photometry (HFP). The isoluminance ratios were used to estimate individual cone spectral sensitivities based on an adapted version of the Individual Colorimetric Observer (ICO) model of Asano, Fairchild, and Blondé [5]. The adapted ICO model described the perceived luminance better than the cone spectral sensitivities of the standard observer. In the second experiment, we measured the temporal contrast sensitivity function for three participants that had substantially different isoluminance ratios. The results give a first impression on the size of individual differences in sensitivity to chromatic temporal modulations. In addition, we proposed two quantitative measures, namely *circularity* and *homogeneity* to describe the local and global uniformity of

a color space, respectively. With such measures, efforts were made to improve both local uniformity and global uniformity of existing color spaces, with a particular focus on the cone excitation spaces LMS and DKL, since they showed to be promising based on the results of **Chapter 3** and **Chapter 4**. Unfortunately, the adapted ICO models did not improve the uniformity of existing color spaces sufficiently (compared to $u'_{10^\circ} v'_{10^\circ}$), which might indicate that more complicated models are needed.

8.1. LIMITATIONS AND GENERAL IMPLICATIONS

With the results summarized above we have addressed the research questions proposed in **Chapter 2**. Unfortunately, we did not find the final answer to all questions with the experiments we performed. This is partly due to the too many unknowns at the start of this research. We here comment on a number of important limitations in the research we did.

8.1.1. CHOICE OF COLOR SPACE

The general goal of this thesis was to develop a temporally uniform color space to ensure high-quality dynamic light. Although we were aware that literature had shown that existing color spaces failed to predict human perception of dynamic colored light transitions, we had to generate our stimuli in one of the existing color spaces. As such, we choose to use Delta-E-per-second to define the speed of dynamic colored light transitions because CIELAB was – and still is – known as a spatially uniform color space, and Delta-E is considered as a good metric to quantify spatial color differences. Our results, however, confirm that CIELAB is not an adequate color space to describe chromatic temporal transitions or modulations. The consequence of nonetheless using CIELAB to generate “linear” chromatic transitions or modulations is that they were not perceptually linear. We expect this effect to be more pronounced for the slow transitions (for which we measured perceived speed) than for the fast transitions (used to measure chromatic flicker), but for both the effect of the particular choice of the color space on the measured results is not known.

For analyzing the results of our psychophysical experiments we also used the cone-excitation spaces LMS and DKL. For transforming the spectral data of the stimuli to LMS values we used the cone spectral sensitivities provided by Stockman and Sharpe [133]. The transformation to DKL values, however, was not unambiguously defined in literature. At a conceptual level, it is widely accepted that L-M represents the red-green opponent channel, while S-(L+M) represents the yellow-blue opponent channel. However, when it comes to computations, opinions are divided on how the cone-excitation signals should be combined. Some literature indicates that the input from the different cones is ‘balanced’ [91, 114, 116, 138], which means that the weights for combining the different cone signals are the same. Other literature states that the ratios for combining the components to the luminance channel are primarily dependent on the individual [131], and that the luminance channel is thus represented as L+1.5M, where the 1.5 is a group average. In the psychophysical experiments we performed, we only corrected for individual isoluminance for fast changes in color (i.e., chromatic flicker), as sensitivity to luminance modulation is higher than sensitivity to chromatic temporal modulation above 4-5Hz [15]. As a result, we were able to derive and adopt the individual weight w (in L+ w M) from the isoluminance ratios we collected. On the contrary, when modeling the slow temporal modulations the individual isoluminance data were not available, and thus, we treated the opponent channels as

balanced (i.e., L-M, S-(L+M)) instead of assigning an arbitrary value. The effect of this choice on our data analysis and modeling in DKL space is still unknown.

8.1.2. ADAPTATION

Adaptation in the visual system is fundamental and ubiquitous [155]. Results of light/dark adaptation [70] and chromatic adaptation [108] have been widely documented. Adaptation in real-life vision is seldom complete, and the degree of adaptation depends on many things, including amongst others the field of view [83]. The visual system is also known to adapt to the dynamics of light, such as flicker. Flicker adaptation has been described in literature [122], but the model for predicting it is still lacking.

Due to the nature of our psychophysical experiments, participants continuously adapted to the luminance, chromaticity and flicker of our stimuli in a largely uncontrolled way. We included attempts to minimize the effect of adaptation during our experiments. For instance, in one of the experiments in **Chapter 3** we showed a ten-second 4000 K constant illumination between the test stimulus and the reference stimulus in order to avoid the effect of adaptation to one stimulus on the perception of the next stimulus. Similarly, in the experiment of **Chapter 5**, we presented an unmodulated two-minute stimulus of a fixed color before the measurement of the detection threshold of chromatic flicker in order to avoid that participants gradually adapted to flicker during the experiment. Nonetheless, uncontrolled effects of adaptation may have affected our experimental results, and attempts to better disentangle adaptation from the measurement of the perceived dynamics of colored light transitions may benefit the development of a temporally uniform color space.

8

8.1.3. INDIVIDUAL DIFFERENCES AND THE USE OF INDIVIDUAL ISOLUMINANCE

Individual differences in visual perception are not surprising and can have several causes. From a physiological point of view, individual differences in vision can be explained by differences in pre-receptoral screening (i.e., related to the density of the lens and macular pigment), differences in photopigment and cone ratios, and post-receptoral differences [94, 154]. Many studies have focused on explaining where individual differences originate from. Though very relevant and of great value in itself, these studies are insufficient to understand how important individual differences are in real applications, such as for designing temporal color transitions that are meant to be pleasant for each individual observer. For that reason we also looked into individual differences in chromatic flicker visibility.

Despite the existence of substantial individual differences, advances in visual perception and its application are often based on the development of models for the standard observer, which is usually determined as a group average of empirical data transformed into a mathematical model. The concept of a standard observer has been used to describe various properties of human vision, such as luminance (i.e., the standard photopic luminous observer, characterized by the luminous efficiency function), chromaticity (i.e., the standard colorimetric observer, characterized by Color Matching Functions, CMFs) and spatial contrast (i.e., the standard spatial observer, defined by [149]). It would be valuable to define a standard temporal observer for describing perceived temporal contrast. This standard observer could then be used to develop a temporal color space.

It is important though to realize that a standard observer is different from any real observer [69], since individuals differ in their perception. The Individual Colorimetric

Observer (ICO) model proposed by Asano, Fairchild, and Blondé [5] has greatly extended the freedom to vary the CIE 2006 Physiological Observer (CIEPO06) [69] for the purpose of looking into individual differences on color perception. To the best of our knowledge, those models on individual differences are mostly limited to the level of photoreceptors. Much less is known on individual differences at the post-photoreceptor processing levels.

The discussion on individual versus averaged observers has multiple entrances in this thesis. Obviously, from a scientific point of view our experimental results are aimed to contribute to the establishment of a standard temporal observer. Yet from an application point of view, one needs to know possible variation around the group average in order to know how to design temporal chromatic transitions that are pleasant to everyone. Therefore, we also looked into individual differences. For defining our stimuli, however, we again made use of standard observer models, for example those on which the CIELAB color space is based. Using this color space, we defined dynamic colored light that changed in chromaticity, while keeping the luminance constant. To do so, we used the standard photopic luminous observer, characterized by the luminous efficiency function $[V(\lambda)]$. Individual observers, however, have a different photopic luminous function, and thus constant luminance does not assure individual isoluminance. So, for our experiments, ideally, all stimuli should be corrected to be individually isoluminant. However, in practice, it is very time-consuming to measure each observer's photopic luminous function. Literature has pointed out that for slower temporal transitions below 4-5 Hz, human subjects are far more sensitive to chromatic modulation than luminance modulation [15]. Thus, for the experiments carried out in **Chapter 3** and **Chapter 4** (with a temporal frequency of 0.33 Hz) and **Chapter 5** (with temporal frequencies of 2 Hz and 4 Hz), we did not implement individual isoluminance corrections. For the experiments carried out in **Chapter 6** and **Chapter 7**, on the other hand, chromatic flicker thresholds were measured for higher frequencies (i.e., between 8 Hz and 25 Hz), and thus individual isoluminance corrections were more relevant, and hence applied. However, instead of measuring the full photopic luminous function, we used a method similar to heterochromatic flicker photometry (HFP). We measured the luminance ratio needed to make a modulation between two extreme colors at a temporal frequency of 25 Hz with a fixed amplitude (i.e., $\Delta(u', v') = 0.05$) perceptually constant. With these results our chromatic flicker stimuli could be designed isoluminant, though in principle only at the measured amplitude of $\Delta(u', v') = 0.05$, and so not necessarily at other amplitudes. The results of a pilot study, in which three amplitudes were measured (i.e., $\Delta(u', v') = 0.01, 0.03$ and 0.05) revealed a linear relation between the amplitude of the chromatic modulation and the logarithm of the corresponding isoluminance ratio. Hence, we adopted a linear interpolation to determine the individual luminance ratio for each specific amplitude. This may, however, have resulted in some inaccuracies.

8.1.4. FIELD OF VIEW, THE INTERACTION OF CENTRAL VISION AND PERIPHERAL VISION

For general applications, temporal dynamics are more realistically represented with a room-size field of view. As a consequence, central and peripheral vision interfere. In **Chapter 3**, the psychophysical experiment was conducted with a room-size field of view. However, since it is still unknown how peripheral and central vision interact in such a field of view, this experimental choice introduced uncertainty. Due to the spatial distribution of rods and cones in the retina, CMFs for a 2-degree and a 10-degree field of view have been developed,

where the former merely represents central vision, and the latter is affected by peripheral vision as well. There is no strict requirement that the stimulus can only have a size of either 2 degrees or 10 degrees. For practical purposes, one uses the 10-degree CMFs if the stimulus covers more than 2 degrees. Obviously, this is an approximation since it does not account for the amount of peripheral vision that contributes to central vision. To reduce that uncertainty, we have limited ourselves to a 10-degree field of view in **Chapter 4**, **Chapter 5** and **Chapter 6**, but this clearly does not guarantee that our results represent perception of colored transitions in full rooms.

Apart from the four major limitations discussed above, we faced additional issues with our experimental studies. Since we are convinced that these remaining points have had less impact on the results of our study, we discuss them less extensively.

8.1.5. LIMITATIONS OF THE HARDWARE

LED-based systems usually provide easy-controllable and accurate colorimetric reproduction of intended colors. However, many LED-based lighting systems also still suffer from color shifts during use as a consequence of temperature changes (i.e., heating). For example, for the system used in **Chapter 5** color shifts could be as large as $0.006 \Delta(u', v')$ over the course of one hour. Such a color shift is perceivable, and as such the colorimetric point presented could be perceived as different from the intended one. To alleviate this issue, we allowed the LED system to properly warm-up before the start of any experimental session. In that case, the color coordinates measured during one hour were cluttered well with color differences of approximately $0.0012 \Delta(u', v')$, which is imperceivable.

8

8.1.6. PSYCHOPHYSICAL EXPERIMENTS

Psychophysics has a history of about 160 years since Fechner [24] proposed the classical psychophysical methods. Since then, hundreds of psychophysical methods have been developed and become common use, such as methods based on the signal detection theory (SDT) or parametric adaptive procedures (e.g., PEST [137]; [153]). Research on psychophysics is still going on, and new, more efficient adaptive procedures are still popping out. For example, QUEST Plus [150] was proposed in 2017 and is considered as an efficient and accurate method. Despite the many choices to be made in psychophysics, the criteria for comparing different methods, being accuracy, precision, and efficiency, remain valid. Evaluating those criteria with newer methods would definitely have added value to the current thesis.

8.2. FUTURE WORK

Our experimental studies are a first step towards developing a uniform temporal color space based on a standard temporal observer, but we are not there yet. The limitations discussed above illustrate the complexity we faced in the process of developing such a color space. As such these limitations also clearly point to directions for further research. The historical development of spatial color spaces in terms of improved uniformity has inspired us in many ways. For example, the MacAdam ellipses were redesigned to become more circular in the CIE 1976 UCS color space than in the CIE 1931 xy chromaticity diagram. In addition, the sizes of the ellipses at different base colors were made more similar by defining CIELAB. Using the same rationale, for a temporal uniform color space, each temporal frequency-dependent ellipse should be more circular and have a consistent shape at different locations in the color

space. We have quantified those two aspects as *circularity* and *homogeneity*, respectively, to describe the local and global uniformity of a color space. Based on the limitations described above we have prioritized some specific directions of research to improve the circularity and homogeneity of a temporal color space.

8.2.1. IMPROVED METHOD FOR ESTIMATING INDIVIDUAL CONE FUNDAMENTALS

In **Chapter 7**, we improved the ICO model, by adding a parameter w to the CIEPO06 chromatic model, where w represents the relative weight between L- and M-cone spectral sensitivities. In the optimization procedure for deriving individual cone fundamentals, we used the isoluminance ratios, by assuming that the S-cone did not contribute to the luminance channel, which is a widely accepted assumption. The attempt to use the ICO model to improve circularity and homogeneity of a LMS-based temporal color space was not a success, probably because individual sensitivities of S-cones could not be derived. Andersen, Finlayson, and Connah [4] proposed to estimate individual cone fundamentals from their color matching functions. We did, however, not choose for this approach, since measuring individual color matching functions is considerably more tedious and time-consuming than using HFP. Although our approach resulted in a better-than-average estimation of the individual cone fundamentals, more research should be done on developing an efficient and accurate method to estimate the individual cone fundamentals. Suppose such a method is developed, the next step is to explore whether the weight between L- and M-cone spectral sensitivities alone is sufficient. If not, individually tuned weights at post-opponent processing may be considered.

8.2.2. IMPROVED COLOR SPACE

In **Chapter 7**, the detection thresholds of chromatic temporal modulations were expressed in different color spaces, i.e., u', v' -chromaticity diagram, the LMS-based color space, and the DKL-based color space, computed with both Stockman-Sharpe (SS) cone fundamentals and ICO-based fundamentals. However, contrary to our expectations, the u', v' -chromaticity diagram has overall the best circularity and homogeneity. The results suggest that different (or more complicated) forms of contrast measures, so not only limited to cone-contrast based measures and DKL-like measures, should be systematically investigated. Besides, as mentioned above, this should be explored in combination with balanced or unbalanced ways of signal coupling between the (LMS or DKL) channels. For example, one could consider the weighting factors in the yellow-blue opponent processing channel $S - (\alpha L + \beta M)$, but that would also require the development of an algorithm to determine the values of α and β .

8.2.3. ADAPTATION

As mentioned earlier, the knowledge on chromatic adaptation to dynamic stimuli is very limited, and definitely insufficient to understand and model the impact of continuous adaptation to the measurement of visibility thresholds for temporal chromatic modulations. In addition, not only chromatic adaptation, but also flicker adaptation may have had some influence on our findings. Therefore, we need additional experiments to measure the degree of adaptation to dynamic chromatic stimuli. These experiments will have to measure the time course of adaptation for different temporal modulations, varying in amplitude and

frequency. The measured speed of adaption will give insight in how big its effect on our findings might have been. If applicable, adaptation related parameters can be added to the temporal color space model to improve its circularity and homogeneity.

8.3. SUMMARY OF CONTRIBUTIONS OF THE CURRENT THESIS AND GENERAL IMPLICATIONS

Our long-term goal is to develop a temporally uniform color space. The main purpose of having such a space is that light designers get means to design temporal colored light transitions with a controllable perceived speed. To address this research goal, several psychophysical experiments were designed to collect data on: 1) the perceived speed of temporal color transitions around different base color points in different directions; and 2) the detection threshold of chromatic flicker at different base color points, modulation directions and temporal frequencies. Those data confirmed that existing color spaces were not able to predict the perception of temporal colored light transitions in a uniform way. Therefore, those data were used to adapt existing color spaces towards a temporally uniform color space. Although these adaptations did not yield a sufficiently uniform temporal color space, our modeling efforts have provided valuable insights.

To the best of our knowledge, our dataset on speed perception and flicker visibility of temporally modulated chromatic light is unique in its kind. It is a valuable dataset for the scientific community, involved in developing a temporally uniform color space. It gives insight in the impact of base color point, modulation direction and modulation frequency on the temporal sensitivity, and provides a first estimate of individual differences in this sensitivity. To quantify uniformity of a color space, we defined a mathematical expression for the local uniformity (i.e., *circularity*) and global uniformity (i.e., *homogeneity*). These measures may be used by standardization bodies to compare the performance of different color spaces.

For the community of light designers, despite the unavailability of a temporally uniform color space yet, some general knowledge may be extracted from our experimental findings. For example, colored light transitions that change in hue need much smaller steps (at a fixed time interval) in the yellow and purple regions of a color space than in the red, green, and blue regions, to have the same perceived speed. When it comes to changes along the chroma direction, larger steps are needed in the yellow and purple regions than in the red, green, and blue regions to have the same perceived speed. This knowledge enables lighting designers to control the speed and smoothen the perception of any change in colored light.

Bibliography

- [1] Edward H. Adelson. “The Delayed Rod Afterimage”. In: *Vision Research* 22.10 (Jan. 1982), pp. 1313–1328. DOI: [10.1016/0042-6989\(82\)90144-4](https://doi.org/10.1016/0042-6989(82)90144-4).
- [2] Sarah R. Allred and Maria Olkkonen. “The Effect of Background and Illumination on Color Identification of Real, 3D Objects”. In: *Frontiers in Psychology* 4 (Nov. 2013). DOI: [10.3389/fpsyg.2013.00821](https://doi.org/10.3389/fpsyg.2013.00821).
- [3] M. Alpern. “Rhodopsin Kinetics in the Human Eye”. In: *The Journal of Physiology* 217.2 (Sept. 1971), pp. 447–471. DOI: [10.1113/jphysiol.1971.sp009580](https://doi.org/10.1113/jphysiol.1971.sp009580).
- [4] Casper F. Andersen, Graham D. Finlayson, and David Connah. “Estimating Individual Cone Fundamentals From Their Color-Matching Functions”. In: *Journal of the Optical Society of America A* 33.8 (July 2016), pp. 1579–1588. DOI: [10.1364/josaa.33.001579](https://doi.org/10.1364/josaa.33.001579).
- [5] Yuta Asano, Mark D. Fairchild, and Laurent Blondé. “Individual Colorimetric Observer Model”. In: *PLOS ONE* 11.2 (Feb. 2016). Ed. by Adrian G Dyer, e0145671. DOI: [10.1371/journal.pone.0145671](https://doi.org/10.1371/journal.pone.0145671).
- [6] Beckhoff. *Fast Communication With the DMX Protocol*. URL: https://download.beckhoff.com/download/document/Application_Notes/DK9222-0311-0029.pdf.
- [7] Wout van Bommel. “Dynamic Lighting At Work – Both in Level and Colour”. In: *Proceedings of the 2nd CIE Expert Symposium on Lighting and Health, 7-8 September 2006, Ottawa, Canada*. Commission internationale de l’éclairage, Sept. 2006, pp. 1–7. URL: <http://cie.co.at/publications/proceedings-2nd-cie-expert-symposium-lighting-and-health-7-8-september-2006-ottawa>.
- [8] PR Boyce et al. “The impact of spectral power distribution on the performance of an achromatic visual task”. In: *Lighting Research & Technology* 35.2 (June 2003), pp. 141–156. DOI: [10.1191/1477153503li0750a](https://doi.org/10.1191/1477153503li0750a).
- [9] D. H. Brainard and L. T. Maloney. “Surface Color Perception and Equivalent Illumination Models”. In: *Journal of Vision* 11.5 (May 2011). DOI: [10.1167/11.5.1](https://doi.org/10.1167/11.5.1).
- [10] G. S. Brindley. “Two New Properties of Foveal After-Images and a Photochemical Hypothesis to Explain Them”. In: *The Journal of Physiology* 164.1 (Oct. 1962), pp. 168–179. DOI: [10.1113/jphysiol.1962.sp007011](https://doi.org/10.1113/jphysiol.1962.sp007011).
- [11] Christine A. Curcio et al. “Distribution and Morphology of Human Cone Photoreceptors Stained With Anti-Blue Opsin”. In: *The Journal of Comparative Neurology* 312.4 (Oct. 1991), pp. 610–624. DOI: [10.1002/cne.903120411](https://doi.org/10.1002/cne.903120411).

- [12] A M Derrington, J Krauskopf, and P Lennie. “Chromatic Mechanisms in Lateral Geniculate Nucleus of Macaque”. In: *The Journal of Physiology* 357.1 (Dec. 1984), pp. 241–265. DOI: [10.1113/jphysiol.1984.sp015499](https://doi.org/10.1113/jphysiol.1984.sp015499).
- [13] Maria Amparo Díez-Ajenjo and Pascual Capilla. “Spatio-temporal Contrast Sensitivity in the Cardinal Directions of the Colour Space. A Review”. In: *Journal of Optometry* 3.1 (2010), pp. 2–19. DOI: [10.3921/joptom.2010.2](https://doi.org/10.3921/joptom.2010.2).
- [14] W. J. Dixon and A. M. Mood. “A Method for Obtaining and Analyzing Sensitivity Data”. In: *Journal of the American Statistical Association* 43.241 (Mar. 1948), pp. 109–126. DOI: [10.1080/01621459.1948.10483254](https://doi.org/10.1080/01621459.1948.10483254).
- [15] Karen R Dobkins, Christina M Anderson, and Barry Lia. “Infant Temporal Contrast Sensitivity Functions (tCSFs) Mature Earlier for Luminance Than for Chromatic Stimuli: Evidence for Precocious Magnocellular Development?” In: *Vision Research* 39.19 (Oct. 1999), pp. 3223–3239. DOI: [10.1016/s0042-6989\(99\)00020-6](https://doi.org/10.1016/s0042-6989(99)00020-6).
- [16] Karen R. Dobkins, Barry Lia, and Davida Y. Teller. “Infant Color Vision: Temporal Contrast Sensitivity Functions for Chromatic (Red/Green) Stimuli in 3-Month-Olds”. In: *Vision Research* 37.19 (Oct. 1997), pp. 2699–2716. DOI: [10.1016/s0042-6989\(97\)81180-7](https://doi.org/10.1016/s0042-6989(97)81180-7).
- [17] Robert F Dougherty, William A Press, and Brian A Wandell. “Perceived Speed of Colored Stimuli”. In: *Neuron* 24.4 (Dec. 1999), pp. 893–899. DOI: [10.1016/s0896-6273\(00\)81036-3](https://doi.org/10.1016/s0896-6273(00)81036-3).
- [18] Walter H. Ehrenstein and Addie Ehrenstein. “Psychophysical Methods”. In: *Modern Techniques in Neuroscience Research*. Springer Berlin Heidelberg, 1999, pp. 1211–1241. DOI: [10.1007/978-3-642-58552-4_43](https://doi.org/10.1007/978-3-642-58552-4_43).
- [19] Auria Eisen-Enosh et al. “Evaluation of Critical Flicker-Fusion Frequency Measurement Methods for the Investigation of Visual Temporal Resolution”. In: *Scientific Reports* 7.1 (Nov. 2017). DOI: [10.1038/s41598-017-15034-z](https://doi.org/10.1038/s41598-017-15034-z).
- [20] Rhea T. Eskew, Charles F. Stromeyer, and Richard E. Kronauer. “Temporal Properties of the Red-Green Chromatic Mechanism”. In: *Vision Research* 34.23 (Dec. 1994), pp. 3127–3137. DOI: [10.1016/0042-6989\(94\)90078-7](https://doi.org/10.1016/0042-6989(94)90078-7).
- [21] Mark D. Fairchild. *Color Appearance Models (3rd Edition)*. John Wiley Sons, Ltd, 2013. ISBN: 9781119967033.
- [22] Mark D. Fairchild and Lisa Reniff. “Time Course of Chromatic Adaptation for Color-Appearance Judgments”. In: *Journal of the Optical Society of America A* 12.5 (May 1995), pp. 824–833. DOI: [10.1364/josaa.12.000824](https://doi.org/10.1364/josaa.12.000824).
- [23] J. C. Falmagne. “Psychophysical Measurement and Theory. Handbook of Perception and Human Performance / Vol. 1, Sensory Processes and Perception.” In: ed. by Kenneth R Boff, Lloyd Kaufman, and James P Thomas. New York: Wiley, 1986. Chap. 1.
- [24] Gustav Theodor Fechner. *Elemente der psychophysik*. Breitkopf u. Härtel, 1860.
- [25] Christine Fernandez-Maloigne and Alain Trémeau. “Color Appearance Models”. In: *Digital Color*. John Wiley & Sons, Inc., m 2013, pp. 65–92. DOI: [10.1002/9781118562680.ch3](https://doi.org/10.1002/9781118562680.ch3).

- [26] Miguel A. García-Pérez. “Forced-Choice Staircases With Fixed Step Sizes: Asymptotic and Small-Sample Properties”. In: *Vision Research* 38.12 (June 1998), pp. 1861–1881. DOI: [10.1016/s0042-6989\(97\)00340-4](https://doi.org/10.1016/s0042-6989(97)00340-4).
- [27] Miguel A. García-Pérez. “Optimal Setups for Forced-Choice Staircases With Fixed Step Sizes”. In: *Spatial Vision* 13.4 (2000), pp. 431–448. DOI: [10.1163/156856800741306](https://doi.org/10.1163/156856800741306).
- [28] Miguel A. García-Pérez. “Properties of Some Variants of Adaptive Staircases With Fixed Step Sizes”. In: *Spatial Vision* 15.3 (2002), pp. 303–321. DOI: [10.1163/15685680260174056](https://doi.org/10.1163/15685680260174056).
- [29] Pedro A. Garcia et al. “Measurement of the Relationship Between Perceived and Computed Color Differences”. In: *Journal of the Optical Society of America A* 24.7 (June 2007), pp. 1823–1829. DOI: [10.1364/josaa.24.001823](https://doi.org/10.1364/josaa.24.001823).
- [30] Miguel A. Garcia-Pérez. “A Cautionary Note on the Use of the Adaptive Up-Down Method”. In: *The Journal of the Acoustical Society of America* 130.4 (Oct. 2011), pp. 2098–2107. DOI: [10.1121/1.3628334](https://doi.org/10.1121/1.3628334).
- [31] Miguel A. Garcia-Pérez. “Yes-No Staircases with Fixed Step Sizes: Psychometric Properties and Optimal Setup”. In: *Optometry and Vision Science* 78.1 (Jan. 2001), pp. 56–64. DOI: [10.1097/00006324-200101010-00015](https://doi.org/10.1097/00006324-200101010-00015).
- [32] Karl R. Gegenfurtner and Robert Ennis. “Fundamentals of Color Vision II: Higher-Order Color Processing”. In: *Handbook of Color Psychology*. Ed. by Andrew J. Elliot, Mark D. Fairchild, and Anna Franklin. Cambridge University Press, 2016, pp. 70–109. DOI: [10.1017/cbo9781107337930.005](https://doi.org/10.1017/cbo9781107337930.005).
- [33] Karl R. Gegenfurtner and Daniel C. Kiper. “Color Vision”. In: *Annual Review of Neuroscience* 26.1 (Mar. 2003), pp. 181–206. DOI: [10.1146/annurev.neuro.26.041002.131116](https://doi.org/10.1146/annurev.neuro.26.041002.131116).
- [34] E. M. Granger and J. C. Heurtley. “Visual Chromaticity-Modulation Transfer Function”. In: *Journal of the Optical Society of America* 63.9 (Sept. 1973), pp. 1173–1174. DOI: [10.1364/josa.63.001173](https://doi.org/10.1364/josa.63.001173).
- [35] H. Grassmann. “Zur Theorie der Farbenmischung”. In: *Annalen der Physik und Chemie* 165.5 (1853), pp. 69–84. DOI: [10.1002/andp.18531650505](https://doi.org/10.1002/andp.18531650505).
- [36] Daniel G. Green. “The Contrast Sensitivity of the Colour Mechanisms of the Human Eye”. In: *The Journal of Physiology* 196.2 (May 1968), pp. 415–429. DOI: [10.1113/jphysiol.1968.sp008515](https://doi.org/10.1113/jphysiol.1968.sp008515).
- [37] J. Guild. “The Colorimetric Properties of the Spectrum”. In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 230.681-693 (June 1931), pp. 149–187. DOI: [10.1098/rsta.1932.0005](https://doi.org/10.1098/rsta.1932.0005).
- [38] Karen L Gunther and Karen R Dobkins. “Individual Differences in Chromatic (Red/Green) Contrast Sensitivity Are Constrained by the Relative Number of L- Versus M-Cones in the Eye”. In: *Vision Research* 42.11 (May 2002), pp. 1367–1378. DOI: [10.1016/s0042-6989\(02\)00043-3](https://doi.org/10.1016/s0042-6989(02)00043-3).
- [39] T. Hansen, L. Pracejus, and K. R. Gegenfurtner. “Color Perception in the Intermediate Periphery of the Visual Field”. In: *Journal of Vision* 9.4 (Apr. 2009), pp. 26–26. DOI: [10.1167/9.4.26](https://doi.org/10.1167/9.4.26).

- [40] Marjolein Hartog. “Dynamic Coloured Lighting: An Interaction Concept for the Creation and Control of Atmospheres at Home”. MA thesis. Delft University of Technology, June 2010. URL: <http://resolver.tudelft.nl/uuid:acf2e118-2b0d-475e-8058-757d20773db8>.
- [41] Lewis O. Harvey. *Detection Theory: Sensory and Decision Processes*. 2010. URL: http://psych.colorado.edu/~lharvey/P4165/P4165_2010_3_Fall/2010_Fall_pdf_Class_Handouts/Detection%20Theory.pdf.
- [42] Markku Hauta-Kasari. *Munsell Colors Matt (Spectrofotometer Measured)*. URL: <https://sites.uefi.fi/spectral/munsell-colors-matt-spectrofotometer-measured/>.
- [43] Georg Hoffmann et al. “Effects of Variable Lighting Intensities and Colour Temperatures on Sulphatoxymelatonin and Subjective Mood in an Experimental Office Workplace”. In: *Applied Ergonomics* 39.6 (Nov. 2008), pp. 719–728. DOI: [10.1016/j.apergo.2007.11.005](https://doi.org/10.1016/j.apergo.2007.11.005).
- [44] Gerard J. C. van der Horst. “Chromatic Flicker*”. In: *Journal of the Optical Society of America* 59.9 (Sept. 1969), pp. 1213–1217. DOI: [10.1364/josa.59.001213](https://doi.org/10.1364/josa.59.001213).
- [45] Gerard J. C. van der Horst and Maarten A. Bouman. “Spatiotemporal Chromaticity Discrimination*”. In: *Journal of the Optical Society of America* 59.11 (Nov. 1969), pp. 1482–1488. DOI: [10.1364/josa.59.001482](https://doi.org/10.1364/josa.59.001482).
- [46] Rong-Hwa Huang et al. “Effects of Correlated Color Temperature on Focused and Sustained Attention Under White LED Desk Lighting”. In: *Color Research & Application* 40.3 (Apr. 2014), pp. 281–286. DOI: [10.1002/co1.21885](https://doi.org/10.1002/co1.21885).
- [47] Cord Huchzermeyer and Jan Kremers. “Perifoveal L- and M-Cone-Driven Temporal Contrast Sensitivities at Different Retinal Illuminances”. In: *Journal of the Optical Society of America A* 33.10 (Sept. 2016), pp. 1989–1998. DOI: [10.1364/josaa.33.001989](https://doi.org/10.1364/josaa.33.001989).
- [48] International Commission on Illumination (CIE). *CIE S 017:2020 ILV: International Lighting Vocabulary (2nd Edition)*. Tech. rep. URL: <https://cie.co.at/eilvterm/17-22-092>.
- [49] John R. Jarvis et al. “Measuring and Modelling the Spatial Contrast Sensitivity of the Chicken (*Gallus G. Domesticus*)”. In: *Vision Research* 49.11 (June 2009), pp. 1448–1454. DOI: [10.1016/j.visres.2009.02.019](https://doi.org/10.1016/j.visres.2009.02.019).
- [50] Amanda Johansson and Monica Sandström. *Sensitivity of the Human Visual System to Amplitude Modulated Light*. Tech. rep. Umeå University, 2003. URL: http://nile.lub.lu.se/arbarch/arb/2003/arb2003_04.pdf.
- [51] Christian Kaernbach. “Simple Adaptive Testing With the Weighted up-Down Method”. In: *Perception & Psychophysics* 49.3 (May 1991), pp. 227–229. DOI: [10.3758/bf03214307](https://doi.org/10.3758/bf03214307).
- [52] Peter K. Kaiser. “Sensation Luminance: A New Name to Distinguish CIE Luminance From Luminance Dependent on an Individual’s Spectral Sensitivity”. In: *Vision Research* 28.3 (Jan. 1988), pp. 455–456. DOI: [10.1016/0042-6989\(88\)90186-1](https://doi.org/10.1016/0042-6989(88)90186-1).
- [53] Peter K. Kaiser and Robert M. Boynton. *Human Color Vision (2nd Edition)*. Optical Society of America, May 1996. ISBN: 1557524610.

- [54] D. H. Kelly. “Spatiotemporal Variation of Chromatic and Achromatic Contrast Thresholds”. In: *Journal of the Optical Society of America* 73.6 (June 1983), pp. 742–750. DOI: [10.1364/josa.73.000742](https://doi.org/10.1364/josa.73.000742).
- [55] D. H. Kelly and D. van Norren. “Two-Band Model of Heterochromatic Flicker”. In: *Journal of the Optical Society of America* 67.8 (Aug. 1977), p. 1081. DOI: [10.1364/josa.67.001081](https://doi.org/10.1364/josa.67.001081).
- [56] P. Khademagha et al. “Implementing Non-Image-Forming Effects of Light in the Built Environment: A Review on What We Need”. In: *Building and Environment* 108 (Nov. 2016), pp. 263–272. DOI: [10.1016/j.buildenv.2016.08.035](https://doi.org/10.1016/j.buildenv.2016.08.035).
- [57] Caitlin Williams Kiley and W. Martin Usrey. “Cortical Processing of Visual Signals”. In: *Neuroscience in the 21st Century*. Springer New York, 2013, pp. 655–674. DOI: [10.1007/978-1-4614-1997-6_24](https://doi.org/10.1007/978-1-4614-1997-6_24).
- [58] Holger Knau. “Thresholds for Detecting Slowly Changing Ganzfeld Luminances”. In: *Journal of the Optical Society of America A* 17.8 (Aug. 2000), pp. 1382–1387. DOI: [10.1364/josaa.17.001382](https://doi.org/10.1364/josaa.17.001382).
- [59] Xiangzhen Kong et al. “Assessing the Temporal Uniformity of CIELAB Hue Angle”. In: *Journal of the Optical Society of America A* 37.4 (Mar. 2020), pp. 521–528. DOI: [10.1364/josaa.384393](https://doi.org/10.1364/josaa.384393).
- [60] Xiangzhen Kong et al. “Modelling Contrast Sensitivity for Chromatic Temporal Modulations”. In: *26th Color and Imaging Conference Final Program and Proceedings*. Color and Imaging Conference. Society for Imaging Science and Technology, 2018, pp. 324–329. DOI: <https://doi.org/10.2352/ISSN.2169-2629.2018.26.324>.
- [61] Xiangzhen Kong et al. “Perceived Speed of Changing Color in Chroma and Hue Directions in CIELAB”. In: *Journal of the Optical Society of America A* 36.6 (May 2019), p. 1022. DOI: [10.1364/josaa.36.001022](https://doi.org/10.1364/josaa.36.001022).
- [62] Leonid L. Kontsevich and Christopher W. Tyler. “Bayesian Adaptive Estimation of Psychometric Slope and Threshold”. In: *Vision Research* 39.16 (Aug. 1999), pp. 2729–2737. DOI: [10.1016/s0042-6989\(98\)00285-5](https://doi.org/10.1016/s0042-6989(98)00285-5).
- [63] Terry K. Koo and Mae Y. Li. “A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research”. In: *Journal of Chiropractic Medicine* 15.2 (June 2016), pp. 155–163. DOI: [10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012).
- [64] J. Kremers, P. K. Kaiser, and B. B. Lee. “Sensitivity of Macaque Retinal Ganglion Cells and Human Observers to Combined Luminance and Chromatic Temporal Modulation”. In: *Journal of the Optical Society of America A* 9.9 (Sept. 1992), pp. 1477–1485. DOI: [10.1364/josaa.9.001477](https://doi.org/10.1364/josaa.9.001477).
- [65] Arie Andries Kruithof. “Tubular Luminescence Lamps for General Illumination”. In: *Philips Technical Review* 6.3 (Mar. 1941), pp. 65–96.
- [66] M. Krystek. “An Algorithm to Calculate Correlated Colour Temperature”. In: *Color Research & Application* 10.1 (1985), pp. 38–40. DOI: [10.1002/col.5080100109](https://doi.org/10.1002/col.5080100109).
- [67] Rolf G. Kuehni. “Hue Uniformity and the Cielab Space and Color Difference Formula”. In: *Color Research & Application* 23.5 (Oct. 1998), pp. 314–322. DOI: [10.1002/\(sici\)1520-6378\(199810\)23:5<314::aid-col7>3.0.co;2-z](https://doi.org/10.1002/(sici)1520-6378(199810)23:5<314::aid-col7>3.0.co;2-z).

- [68] Rikard Küller et al. “The Impact of Light and Colour on Psychological Mood: A Cross-Cultural Study of Indoor Work Environments”. In: *Ergonomics* 49.14 (Nov. 2006), pp. 1496–1507. DOI: [10.1080/00140130600858142](https://doi.org/10.1080/00140130600858142).
- [69] Commission Internationale de L’Eclairage (CIE). *Fundamental Chromaticity Diagram With Physiological Axes*. Vienna, Austria: Commission internationale de l’éclairage, 2006. ISBN: 9783901906466. DOI: <https://cie.co.at/publications/fundamental-chromaticity-diagram-physiological-axes-part-1>.
- [70] T.D. Lamb and E.N. Pugh. “Dark Adaptation and the Retinoid Cycle of Vision”. In: *Progress in Retinal and Eye Research* 23.3 (May 2004), pp. 307–380. DOI: [10.1016/j.preteyeres.2004.03.001](https://doi.org/10.1016/j.preteyeres.2004.03.001).
- [71] H. de Lange Dzn. “Research into the Dynamic Nature of the Human Fovea→Cortex Systems with Intermittent and Modulated Light I Attenuation Characteristics with White and Colored Light”. In: *Journal of the Optical Society of America* 48.11 (Nov. 1958), pp. 777–784. DOI: [10.1364/josa.48.000777](https://doi.org/10.1364/josa.48.000777).
- [72] B B Lee, P R Martin, and A Valberg. “The Physiological Basis of Heterochromatic Flicker Photometry Demonstrated in the Ganglion Cells of the Macaque Retina.” In: *The Journal of Physiology* 404.1 (Oct. 1988), pp. 323–347. DOI: [10.1113/jphysiol.1988.sp017292](https://doi.org/10.1113/jphysiol.1988.sp017292).
- [73] B. B. Lee, H. Sun, and A. Valberg. “Segregation of Chromatic and Luminance Signals Using a Novel Grating Stimulus”. In: *The Journal of Physiology* 589.1 (Oct. 2010), pp. 59–73. DOI: [10.1113/jphysiol.2010.188862](https://doi.org/10.1113/jphysiol.2010.188862).
- [74] Cassandra R. Lee et al. “Predicting Color Matches From Luminance Matches”. In: *Journal of the Optical Society of America A* 37.4 (Feb. 2020), A35–A43. DOI: [10.1364/josaa.381256](https://doi.org/10.1364/josaa.381256).
- [75] Marjorie R. Leek. “Adaptive Procedures in Psychophysical Research”. In: *Perception & Psychophysics* 63.8 (Nov. 2001), pp. 1279–1292. DOI: [10.3758/bf03194543](https://doi.org/10.3758/bf03194543).
- [76] Moshe Levison and Frank Restle. “Invalid Results From the Method of Constant Stimuli”. In: *Perception & Psychophysics* 4.2 (Mar. 1968), pp. 121–122. DOI: [10.3758/bf03209522](https://doi.org/10.3758/bf03209522).
- [77] H. Levitt. “Transformed Up-Down Methods in Psychoacoustics”. In: *The Journal of the Acoustical Society of America* 49.2B (Feb. 1971), pp. 467–477. DOI: [10.1121/1.1912375](https://doi.org/10.1121/1.1912375).
- [78] Christophe Leys et al. “Detecting Outliers: Do Not Use Standard Deviation Around the Mean, Use Absolute Deviation Around the Median”. In: *Journal of Experimental Social Psychology* 49.4 (July 2013), pp. 764–766. DOI: [10.1016/j.jesp.2013.03.013](https://doi.org/10.1016/j.jesp.2013.03.013).
- [79] B Li et al. “Atmosphere Perception of Dynamic LED Lighting Over Different Hue Ranges”. In: *Lighting Research & Technology* 51.5 (Apr. 2017), pp. 682–703. DOI: [10.1177/1477153517702532](https://doi.org/10.1177/1477153517702532).
- [80] Alexander D. Logvinenko. “Individual Differences in Human Colour Vision as Derived From Stiles & Burch 10° Colour Matching Functions”. In: *Color Research & Application* 38.2 (Jan. 2012), pp. 96–109. DOI: [10.1002/col.20741](https://doi.org/10.1002/col.20741).

- [81] Ronnier Luo, ed. *Encyclopedia of Color Science and Technology*. Springer-Verlag New York, 2016. ISBN: 9781441980724.
- [82] S. M. Luria and David F. Neri. “Individual Differences in Luminous Efficiency Measured by Flicker Photometry”. In: *Color Research & Application* 11.1 (1986), pp. 72–75. DOI: [10.1002/col.5080110113](https://doi.org/10.1002/col.5080110113).
- [83] Shining Ma et al. “Effect of Adapting Field Size on Chromatic Adaptation”. In: *Optics Express* 28.12 (May 2020), pp. 17266–17285. DOI: [10.1364/oe.392844](https://doi.org/10.1364/oe.392844).
- [84] David L. MacAdam. “Visual Sensitivities to Color Differences in Daylight*”. In: *Journal of the Optical Society of America* 32.5 (May 1942), pp. 247–274. DOI: [10.1364/josa.32.000247](https://doi.org/10.1364/josa.32.000247).
- [85] Donald I. A. MacLeod and Robert M. Boynton. “Chromaticity Diagram Showing Cone Excitation by Stimuli of Equal Luminance”. In: *Journal of the Optical Society of America* 69.8 (Aug. 1979), pp. 1183–1186. DOI: [10.1364/josa.69.001183](https://doi.org/10.1364/josa.69.001183).
- [86] Jasna Martinovic. “Magno-, Parvo-, Koniocellular Pathways”. In: *Encyclopedia of Color Science and Technology*. Springer Berlin Heidelberg, 2015, pp. 1–5. DOI: [10.1007/978-3-642-27851-8_278-1](https://doi.org/10.1007/978-3-642-27851-8_278-1).
- [87] S. C. Masin, V. Fanton, and L. Crestoni. “An Experimental Study of the Asymmetry Effect in the Method of Constant Stimuli”. In: *Psychological Research* 50.3 (Dec. 1988), pp. 181–182. DOI: [10.1007/bf00310179](https://doi.org/10.1007/bf00310179).
- [88] George Mather. *Foundations of Sensation and Perception (3rd Edition)*. Psychology Press, 2016. ISBN: 9781848723443.
- [89] D.J. McKeefry, I.J. Murray, and J.J. Kulikowski. “Red–Green and Blue–Yellow Mechanisms Are Matched in Sensitivity for Temporal and Spatial Modulation”. In: *Vision Research* 41.2 (Jan. 2001), pp. 245–255. DOI: [10.1016/s0042-6989\(00\)00247-9](https://doi.org/10.1016/s0042-6989(00)00247-9).
- [90] D.J. McKeefry and M.P. Burton. “The Perception of Speed Based on L-M and S-(LM) Cone Opponent Processing”. In: *Vision Research* 49.8 (May 2009), pp. 870–876. DOI: [10.1016/j.visres.2009.03.004](https://doi.org/10.1016/j.visres.2009.03.004).
- [91] Declan J. McKeefry, Neil R. A. Parry, and Ian J. Murray. “Simple Reaction Times in Color Space: The Influence of Chromaticity, Contrast, and Cone Opponency”. In: *Investigative Ophthalmology & Visual Science* 44.5 (May 2003), pp. 2267–2276. DOI: [10.1167/iovs.02-0772](https://doi.org/10.1167/iovs.02-0772).
- [92] Qianli Meng et al. “Age-Related Changes in Local and Global Visual Perception”. In: *Journal of Vision* 19.1 (Jan. 2019), pp. 1–12. DOI: [10.1167/19.1.10](https://doi.org/10.1167/19.1.10).
- [93] Andrew B. Metha and Kathy T. Mullen. “Temporal Mechanisms Underlying Flicker Detection and Identification for Red–Green and Achromatic Stimuli”. In: *Journal of the Optical Society of America A* 13.10 (Oct. 1996), pp. 1969–1980. DOI: [10.1364/josaa.13.001969](https://doi.org/10.1364/josaa.13.001969).
- [94] John D. Mollon et al. “Individual Differences in Visual Science: What Can Be Learned and What Is Good Experimental Practice?” In: *Vision Research* 141 (Dec. 2017), pp. 4–15. DOI: [10.1016/j.visres.2017.11.001](https://doi.org/10.1016/j.visres.2017.11.001).
- [95] Michael S. Mott et al. “Illuminating the Effects of Dynamic Lighting on Student Learning”. In: *SAGE Open* 2.2 (Apr. 2012), pp. 1–9. DOI: [10.1177/2158244012445585](https://doi.org/10.1177/2158244012445585).

- [96] Michael J. Murdoch. “Characterization and Control of a Multi-Primary LED Light Lab”. In: *Optics Express* 25.24 (Nov. 2017), pp. 29605–29616. DOI: [10.1364/oe.25.029605](https://doi.org/10.1364/oe.25.029605).
- [97] Michael J. Murdoch, Dragan Sekulovski, and Pieter Seuntjens. “The Influence of Speed and Amplitude on Visibility and Perceived Subtlety of Dynamic Light”. In: *19th Color and Imaging Conference Final Program and Proceedings*. Color and Imaging Conference. Society for Imaging Science and Technology, 2011, pp. 265–269.
- [98] Debarshi Mustafi, Andreas H. Engel, and Krzysztof Palczewski. “Structure of Cone Photoreceptors”. In: *Progress in Retinal and Eye Research* 28.4 (July 2009), pp. 289–302. DOI: [10.1016/j.preteyeres.2009.05.003](https://doi.org/10.1016/j.preteyeres.2009.05.003).
- [99] TA Nealey and JH Maunsell. “Magnocellular and Parvocellular Contributions to the Responses of Neurons in Macaque Striate Cortex”. In: *The Journal of Neuroscience* 14.4 (Apr. 1994), pp. 2069–2079. DOI: [10.1523/jneurosci.14-04-02069.1994](https://doi.org/10.1523/jneurosci.14-04-02069.1994).
- [100] John A. Nevin. “Signal Detection Theory and Operant Behavior: A Review of David M. Green and John A. Swets’ Signal Detection Theory and Psychophysics”. In: *Journal of the Experimental Analysis of Behavior* 12.3 (May 1969), pp. 475–480. DOI: [10.1901/jeab.1969.12-475](https://doi.org/10.1901/jeab.1969.12-475).
- [101] David Nguyen-Tri and Jocelyn Faubert. “The Perceived Speed of Drifting Chromatic Gratings Is Mechanism-Dependent”. In: *Vision Research* 42.17 (Aug. 2002), pp. 2073–2079. DOI: [10.1016/s0042-6989\(02\)00127-x](https://doi.org/10.1016/s0042-6989(02)00127-x).
- [102] Nam-Kyu Park and Cheryl A. Farr. “The Effects of Lighting on Consumers’ Emotions and Behavioral Intentions in a Retail Environment: A Cross-Cultural Comparison”. In: *Journal of Interior Design* 33.1 (Sept. 2007), pp. 17–32. DOI: [10.1111/j.1939-1668.2007.tb00419.x](https://doi.org/10.1111/j.1939-1668.2007.tb00419.x).
- [103] Denis G. Pelli and Peter Bex. “Measuring Contrast Sensitivity”. In: *Vision Research* 90 (Sept. 2013), pp. 10–14. DOI: [10.1016/j.visres.2013.04.015](https://doi.org/10.1016/j.visres.2013.04.015).
- [104] Denis G. Pelli and Bart Farell. *Psychophysical Methods. Handbook of Optics: Volume I - Fundamentals, Techniques, and Design (2nd Edition)*. Ed. by Michael Bass. The McGraw-Hill Companies, Inc., 1995. URL: [http://inis.jinr.ru/sl/P_Physics/PE_Electromagnetism/PEo_Optics/Bass%20M.,%20et%20al.%20\(eds.\)%2000SA%20Handbook%20of%20Optics,%20vol.%201%20\(MGH,%201995\)\(1606s\).pdf](http://inis.jinr.ru/sl/P_Physics/PE_Electromagnetism/PEo_Optics/Bass%20M.,%20et%20al.%20(eds.)%2000SA%20Handbook%20of%20Optics,%20vol.%201%20(MGH,%201995)(1606s).pdf).
- [105] M. Perz, D. Sekulovski, and M. Murdoch. “Chromatic Flicker Perception in Human Peripheral Vision Under Mental Load”. In: *18th Color and Imaging Conference Final Program and Proceedings*. Color and Imaging Conference. Society for Imaging Science and Technology, 2010, pp. 33–37. URL: <https://www.ingentaconnect.com/content/ist/cic/2010/00002010/00000001/art00007>.
- [106] Anja Podlesek and Luka Komidar. “Comparison of Three Psychophysical Methods for Measuring Displacement in Frontal Plane Motion”. In: *Review of Psychology* 13.1 (2006), pp. 51–60.
- [107] Thomas H. Rammsayer. “An Experimental Comparison of the Weighted up-Down Method and the Transformed up-Down Method”. In: *Bulletin of the Psychonomic Society* 30.5 (Nov. 1992), pp. 425–427. DOI: [10.3758/bf03334107](https://doi.org/10.3758/bf03334107).

- [108] Oliver Rinner and Karl R Gegenfurtner. “Time Course of Chromatic Adaptation for Color Appearance and Discrimination”. In: *Vision Research* 40.14 (June 2000), pp. 1813–1826. DOI: [10.1016/s0042-6989\(00\)00050-x](https://doi.org/10.1016/s0042-6989(00)00050-x).
- [109] C. Ripamonti et al. “The S-Cone Contribution to Luminance Depends on the M- and L-Cone Adaptation Levels: Silent Surrounds?” In: *Journal of Vision* 9.3 (Mar. 2009), pp. 1–16. DOI: [10.1167/9.3.10](https://doi.org/10.1167/9.3.10).
- [110] J. G. Robson. “Spatial and Temporal Contrast-Sensitivity Functions of the Visual System”. In: *Journal of the Optical Society of America* 56.8 (Aug. 1966), pp. 1141–1142. DOI: [10.1364/josa.56.001141](https://doi.org/10.1364/josa.56.001141).
- [111] Jyrki M. Rovamo, Mia I. Kankaanpää, and Heljä Kukkonen. “Modelling Spatial Contrast Sensitivity Functions for Chromatic and Luminance-Modulated Gratings”. In: *Vision Research* 39.14 (June 1999), pp. 2387–2398. DOI: [10.1016/s0042-6989\(98\)00273-9](https://doi.org/10.1016/s0042-6989(98)00273-9).
- [112] Muhammad Safdar et al. “Perceptually Uniform Color Space for Image Signals Including High Dynamic Range and Wide Gamut”. In: *Optics Express* 25.13 (June 2017), pp. 15131–15151. DOI: [10.1364/oe.25.015131](https://doi.org/10.1364/oe.25.015131).
- [113] Ken Sagawa and Yasuro Takahashi. “Spectral Luminous Efficiency as a Function of Age”. In: *Journal of the Optical Society of America A* 18.11 (Nov. 2001), pp. 2659–2667. DOI: [10.1364/josaa.18.002659](https://doi.org/10.1364/josaa.18.002659).
- [114] Masato Sakurai and Kathy T. Mullen. “Cone Weights for the Two Cone-Opponent Systems in Peripheral Vision and Asymmetries of Cone Contrast Sensitivity”. In: *Vision Research* 46.26 (Dec. 2006), pp. 4346–4354. DOI: [10.1016/j.visres.2006.08.016](https://doi.org/10.1016/j.visres.2006.08.016).
- [115] B. A. Salters and M. P. C. M. Krijn. “Color Reproduction for LED-Based General Lighting”. In: *Nonimaging Optics and Efficient Illumination Systems III*. Ed. by Roland Winston and Pablo Benitez. SPIE, Aug. 2006. DOI: [10.1117/12.680515](https://doi.org/10.1117/12.680515).
- [116] Marcel J. Sankeralli and Kathy T. Mullen. “Estimation of the L-, M-, and S-Cone Weights of the Postreceptoral Detection Mechanisms”. In: *Journal of the Optical Society of America A* 13.5 (May 1996), pp. 906–915. DOI: [10.1364/josaa.13.000906](https://doi.org/10.1364/josaa.13.000906).
- [117] Abhijit Sarkar et al. “From CIE 2006 Physiological Model to Improved Age-Dependent and Average Colorimetric Observers”. In: *Journal of the Optical Society of America A* 28.10 (Sept. 2011), pp. 2033–2048. DOI: [10.1364/josaa.28.002033](https://doi.org/10.1364/josaa.28.002033).
- [118] E. F. Schubert and J. K. Kim. “Solid-State Light Sources Getting Smart”. In: *Science* 308.5726 (May 2005), pp. 1274–1278. DOI: [10.1126/science.1108712](https://doi.org/10.1126/science.1108712).
- [119] Robert Sekuler and Albert Erlebacher. “The Invalidity of “Invalid Results From the Method of Constant Stimuli”: A Common Artifact in the Methods of Psychophysics”. In: *Perception & Psychophysics* 9.3 (May 1971), pp. 309–311. DOI: [10.3758/bf03212655](https://doi.org/10.3758/bf03212655).
- [120] Dragan Sekulovski et al. “Changing Color Over Time”. In: *Ergonomics and Health Aspects of Work with Computers*. Springer Berlin Heidelberg, 2011, pp. 218–225. DOI: [10.1007/978-3-642-21716-6_23](https://doi.org/10.1007/978-3-642-21716-6_23).

- [121] Dragan Sekulovski et al. “Smoothness and Flicker Perception of Temporal Color Transitions”. In: *15th Color Imaging Conference Final Program and Proceedings*. Color and Imaging Conference. Society for Imaging Science and Technology, 2007, pp. 112–117.
- [122] S. Shady, D. I. A. MacLeod, and H. S. Fisher. “Adaptation From Invisible Flicker”. In: *Proceedings of the National Academy of Sciences* 101.14 (Mar. 2004), pp. 5170–5173. DOI: [10.1073/pnas.0303452101](https://doi.org/10.1073/pnas.0303452101).
- [123] Lindsay T. Sharpe et al. “A Luminous Efficiency Function, $V^*(\lambda)$, for Daylight Adaptation”. In: *Journal of Vision* 5.11 (Dec. 2005), pp. 948–968. DOI: [10.1167/5.11.3](https://doi.org/10.1167/5.11.3).
- [124] Steven K. Shevell and Paul R. Martin. “Color Opponency: Tutorial”. In: *Journal of the Optical Society of America A* 34.7 (June 2017), pp. 1099–1108. DOI: [10.1364/josaa.34.001099](https://doi.org/10.1364/josaa.34.001099).
- [125] Martyn Shuttleworth. *Counterbalanced Measures Design*. URL: <https://explorable.com/counterbalanced-measures-design>.
- [126] William A. Simpson. “The Method of Constant Stimuli Is Efficient”. In: *Perception & Psychophysics* 44.5 (Sept. 1988), pp. 433–436. DOI: [10.3758/bf03210427](https://doi.org/10.3758/bf03210427).
- [127] Simon E. Skalicky. “The Lateral Geniculate Nucleus”. In: *Ocular and Visual Physiology*. Springer Singapore, 2016, pp. 201–206. DOI: [10.1007/978-981-287-846-5_13](https://doi.org/10.1007/978-981-287-846-5_13).
- [128] Samuel G. Solomon, Andrew J. R. White, and Paul R. Martin. “Temporal Contrast Sensitivity in the Lateral Geniculate Nucleus of a New World Monkey, the Marmoset *Callithrix jacchus*”. In: *The Journal of Physiology* 517.3 (June 1999), pp. 907–917. DOI: [10.1111/j.1469-7793.1999.0907s.x](https://doi.org/10.1111/j.1469-7793.1999.0907s.x).
- [129] Peep F. M. Stalmeier and Charles M. M. de Weert. “On the Conditions Affecting the Contribution of Color to Brightness Perception”. In: *Color Research & Application* 19.3 (June 1994), pp. 192–201. DOI: [10.1002/col.5080190307](https://doi.org/10.1002/col.5080190307).
- [130] Andrew Stockman and David H. Brainard. *Color Vision Mechanisms. Handbook of Optics: Volume III - Vision and Vision Optics (3rd Edition)*. Ed. by Michael Bass. The McGraw-Hill Companies, Inc., 2010. ISBN: 9780071498913. URL: <https://www.accessengineeringlibrary.com/content/book/9780071498913>.
- [131] Andrew Stockman and David H. Brainard. “Fundamentals of color vision I: color processing in the eye”. In: *Handbook of Color Psychology*. Ed. by Andrew J. Elliot, Mark D. Fairchild, and Anna Franklin. Cambridge University Press, 2016, pp. 27–69. DOI: [10.1017/cbo9781107337930.004](https://doi.org/10.1017/cbo9781107337930.004).
- [132] Andrew Stockman, G. Bruce Henning, and Andrew T. Rider. “Linear–Nonlinear Models of the Red–Green Chromatic Pathway”. In: *Journal of Vision* 17.13 (Nov. 2017), p. 7. DOI: [10.1167/17.13.7](https://doi.org/10.1167/17.13.7).
- [133] Andrew Stockman and Lindsay T. Sharpe. “The Spectral Sensitivities of the Middle- and Long-Wavelength-Sensitive Cones Derived From Measurements in Observers of Known Genotype”. In: *Vision Research* 40.13 (June 2000), pp. 1711–1737. DOI: [10.1016/s0042-6989\(00\)00021-3](https://doi.org/10.1016/s0042-6989(00)00021-3).

- [134] Hao Sun, Joel Pokorny, and Vivianne C. Smith. “Control of the Modulation of Human Photoreceptors”. In: *Color Research Application* 26.S1 (Dec. 2000), S69–S75. DOI: [https://doi.org/10.1002/1520-6378\(2001\)26:1+<::AID-COL16>3.0.CO;2-A](https://doi.org/10.1002/1520-6378(2001)26:1+<::AID-COL16>3.0.CO;2-A). URL: [https://onlinelibrary.wiley.com/doi/abs/10.1002/1520-6378\(2001\)26:1%20%3C::AID-COL16%3E3.0.CO;2-A](https://onlinelibrary.wiley.com/doi/abs/10.1002/1520-6378(2001)26:1%20%3C::AID-COL16%3E3.0.CO;2-A).
- [135] William H. Swanson, Fei Pan, and Barry B. Lee. “Chromatic Temporal Integration and Retinal Eccentricity: Psychophysics, Neurometric Analysis and Cortical Pooling”. In: *Vision Research* 48.26 (Nov. 2008), pp. 2657–2662. DOI: [10.1016/j.visres.2008.03.002](https://doi.org/10.1016/j.visres.2008.03.002).
- [136] William H. Swanson et al. “Temporal Modulation Sensitivity and Pulse-Detection Thresholds for Chromatic and Luminance Perturbations”. In: *Journal of the Optical Society of America A* 4.10 (Oct. 1987), pp. 1992–2005. DOI: [10.1364/josaa.4.001992](https://doi.org/10.1364/josaa.4.001992).
- [137] M. M. Taylor and C. Douglas Creelman. “PEST: Efficient Estimates on Probability Functions”. In: *The Journal of the Acoustical Society of America* 41.4A (Apr. 1967), pp. 782–787. DOI: [10.1121/1.1910407](https://doi.org/10.1121/1.1910407).
- [138] Harald J. Teufel and Christian Wehrhahn. “Evidence for the Contribution of S Cones to the Detection of Flicker Brightness and Red–Green”. In: *Journal of the Optical Society of America A* 17.6 (June 2000), pp. 994–1006. DOI: [10.1364/josaa.17.000994](https://doi.org/10.1364/josaa.17.000994).
- [139] Bernhard Treutwein. “Adaptive Psychophysical Procedures”. In: *Vision Research* 35.17 (Sept. 1995), pp. 2503–2522. DOI: [10.1016/0042-6989\(95\)00016-x](https://doi.org/10.1016/0042-6989(95)00016-x).
- [140] Carroll Vance Truss. “Chromatic Flicker Fusion Frequency as a Function of Chromaticity Difference*”. In: *Journal of the Optical Society of America* 47.12 (Dec. 1957), p. 1130. DOI: [10.1364/josa.47.001130](https://doi.org/10.1364/josa.47.001130).
- [141] Christopher W. Tyler and Andrei Gorea. “Different Encoding Mechanisms for Phase and Contrast”. In: *Vision Research* 26.7 (Jan. 1986), pp. 1073–1082. DOI: [10.1016/0042-6989\(86\)90042-8](https://doi.org/10.1016/0042-6989(86)90042-8).
- [142] J.A. Veitch and G.R. Newsham. “Lighting Quality and Energy-Efficiency Effects on Task Performance, Mood, Health, Satisfaction, and Comfort”. In: *Journal of the Illuminating Engineering Society* 27.1 (Jan. 1998), pp. 107–129. DOI: [10.1080/00994480.1998.10748216](https://doi.org/10.1080/00994480.1998.10748216).
- [143] F Viénot. “Cone Fundamentals: A Model for the Future of Colorimetry”. In: *Lighting Research & Technology* 48.1 (Dec. 2015), pp. 5–13. DOI: [10.1177/1477153515624267](https://doi.org/10.1177/1477153515624267).
- [144] Ingrid Vogels, Dragan Sekulovski, and Bartjan Rijs. “Discrimination and Preference of Temporal Color Transitions”. In: *15th Color Imaging Conference Final Program and Proceedings*. Color and Imaging Conference. Society for Imaging Science and Technology, 2007, pp. 118–121.
- [145] Ingrid Vogels, Dragan Sekulovski, and Bartjan Rijs. “How to Create Appealing Temporal Color Transitions?” In: *Journal of the Society for Information Display* 17.1 (2009), p. 23. DOI: [10.1889/jsid17.1.23](https://doi.org/10.1889/jsid17.1.23).

- [146] Ingrid Vogels, Maartje de Vries, and Thomas van Erp. “Effect of Coloured Light on Atmosphere Perception”. In: *Proceedings of AIC 2008, Colour – Effects and Affects, Interim Meeting of the International Color Association*. Ed. by Karin Fridell Anter Iman Kortbawi Berit Bergström. <https://www.aic-color.org/publications-proceedings>; 2008; Paper No 60. The International Color Association. Stockholm: The Swedish Colour Centre Foundation/ Scandinavian Colour Institute AB, June 2008.
- [147] Gunther Wagner and Robert M. Boynton. “Comparison of Four Methods of Heterochromatic Photometry”. In: *Journal of the Optical Society of America* 62.12 (Dec. 1972), pp. 1508–1515. DOI: [10.1364/josa.62.001508](https://doi.org/10.1364/josa.62.001508).
- [148] Brian A. Wandell. *Foundations of Vision*. Sinauer Associates Inc., 1995. ISBN: 0878938532.
- [149] Andrew B. Watson. “A Spatial Standard Observer for Visual Technology”. In: *Human Vision and Electronic Imaging X*. Ed. by Bernice E. Rogowitz, Thrasyvoulos N. Pappas, and Scott J. Daly. SPIE, Mar. 2005. DOI: [10.1117/12.601935](https://doi.org/10.1117/12.601935).
- [150] Andrew B. Watson. “QUEST: A General Multidimensional Bayesian Adaptive Psychometric Method”. In: *Journal of Vision* 17.3 (Apr. 2017), pp. 1–27. DOI: [10.1167/17.3.10](https://doi.org/10.1167/17.3.10).
- [151] Andrew B. Watson. “Temporal Sensitivity. Handbook of Perception and Human Performance, Vol. 1: Sensory Processes and Perception”. In: ed. by Kenneth R Boff, Lloyd Kaufman, and James P Thomas. New York: Wiley, 1986. Chap. 6.
- [152] Andrew B. Watson and Andrew Fitzhugh. “The Method of Constant Stimuli Is Inefficient”. In: *Perception & Psychophysics* 47.1 (Jan. 1990), pp. 87–91. DOI: [10.3758/bf03208169](https://doi.org/10.3758/bf03208169).
- [153] Andrew B. Watson and Denis G. Pelli. “Quest: A Bayesian Adaptive Psychometric Method”. In: *Perception & Psychophysics* 33.2 (Mar. 1983), pp. 113–120. DOI: [10.3758/bf03202828](https://doi.org/10.3758/bf03202828).
- [154] Michael A. Webster. “Individual Differences in Color Vision”. In: *Handbook of Color Psychology*. Ed. by Andrew J. Elliot, Mark D. Fairchild, and Anna Franklin. Cambridge University Press, 2016, pp. 197–215. DOI: [10.1017/cbo9781107337930.010](https://doi.org/10.1017/cbo9781107337930.010).
- [155] Michael A. Webster. “Visual Adaptation”. In: *Annual Review of Vision Science* 1.1 (Nov. 2015), pp. 547–567. DOI: [10.1146/annurev-vision-082114-035509](https://doi.org/10.1146/annurev-vision-082114-035509).
- [156] Michael A. Webster and J.D. Mollon. “The Influence of Contrast Adaptation on Color Appearance”. In: *Vision Research* 34.15 (Aug. 1994), pp. 1993–2020. DOI: [10.1016/0042-6989\(94\)90028-0](https://doi.org/10.1016/0042-6989(94)90028-0).
- [157] Felix A. Wichmann and N. Jeremy Hill. “The Psychometric Function: II. Bootstrap-Based Confidence Intervals and Sampling”. In: *Perception & Psychophysics* 63.8 (Nov. 2001), pp. 1314–1329. DOI: [10.3758/bf03194545](https://doi.org/10.3758/bf03194545).
- [158] Craig C. Wier, Walt Jesteadt, and David M. Green. “A Comparison of Method-of-Adjustment and Forced-Choice Procedures in Frequency Discrimination”. In: *Perception & Psychophysics* 19.1 (Jan. 1976), pp. 75–79. DOI: [10.3758/bf03199389](https://doi.org/10.3758/bf03199389).
- [159] Billy R. Wooten et al. “A Practical Method of Measuring the Human Temporal Contrast Sensitivity Function”. In: *Biomedical Optics Express* 1.1 (July 2010), p. 47. DOI: [10.1364/boe.1.000047](https://doi.org/10.1364/boe.1.000047).

- [160] W D Wright. “A Re-Determination of the Trichromatic Coefficients of the Spectral Colours”. In: *Transactions of the Optical Society* 30.4 (Mar. 1929), pp. 141–164. DOI: [10.1088/1475-4878/30/4/301](https://doi.org/10.1088/1475-4878/30/4/301).
- [161] Sophie Wuergler and Kaida Xiao. “Color Vision, Opponent Theory”. In: *Encyclopedia of Color Science and Technology*. Springer New York, 2016, pp. 413–418. DOI: [10.1007/978-1-4419-8071-7_92](https://doi.org/10.1007/978-1-4419-8071-7_92).
- [162] Wonyoung Yang and Jin Yong Jeon. “Effects of Correlated Colour Temperature of LED Light on Visual Sensation, Perception, and Cognitive Performance in a Classroom Lighting Environment”. In: *Sustainability* 12.10 (May 2020), p. 4051. DOI: [10.3390/su12104051](https://doi.org/10.3390/su12104051).
- [163] Qasim Zaidi et al. “Neural Locus of Color Afterimages”. In: *Current Biology* 22.3 (Feb. 2012), pp. 220–224. DOI: [10.1016/j.cub.2011.12.021](https://doi.org/10.1016/j.cub.2011.12.021).
- [164] Andrew J. Zele, Vivianne C. Smith, and Joel Pokorny. “Spatial and Temporal Chromatic Contrast: Effects on Chromatic Discrimination for Stimuli Varying In L- and M-Cone Excitation”. In: *Visual Neuroscience* 23.3-4 (May 2006), pp. 495–501. DOI: [10.1017/s0952523806232012](https://doi.org/10.1017/s0952523806232012).

Summary

LED (light-emitting diode) technology has unleashed new options for both general lighting and decorative lighting, partly because of its fast temporal response enabling inexpensive ways to create dynamic colored light. The speed of those dynamics can range from very *fast* (i.e., only the colors at the start and the end of the transition are perceived) to very *slow* (i.e., the transition is hardly noticeable). To make these dynamics attractive to observers, the light should change smoothly over time in luminance and/or chromaticity, and at the desired perceived rate of change. To achieve this, one needs a color space in which temporal color differences can be described uniformly from a perception perspective. As such, this project's long-term goal is to develop a temporal uniform color space to create smooth dynamic light effects with well-defined perceived speed.

Establishing a temporally uniform color space is a task too big for a single PhD-project. However, with this long-term goal in mind, several psychophysical experiments were carried out, addressing specifically one aspect of high-quality dynamics - the perceived speed of a colored light transition. We sought inspiration in the historical development of spatially uniform color spaces and designed a series of psychophysical experiments to approach our goal. Those psychophysical experiments addressed the following research questions: (1) what is the effect of modulation direction and location in the color space on the perceived speed of a colored light transition; (2) what is an efficient yet accurate and precise method for measuring perceived flicker of colored light modulations; (3) how can we model the effect of temporal frequency on chromatic flicker visibility for a wide range of base colors and chromatic modulations; (4) what is a suitable contrast measure to describe temporal contrast sensitivity functions (TCSFs); and (5) can we reduce individual differences in TCSFs by taking into account individual cone spectral sensitivities?

In **Chapter 3**, we compared the perceived speed of periodic temporal transitions in CIELAB chroma and hue directions around five base colors (i.e., the five Munsell hues: 5R (red), 5Y (yellow), 5G (green), 5B (blue), and 5P (purple)). The experiment was conducted with stimuli consisting of full-room illumination. In sequential paired presentations, observers were asked to identify which transition appeared faster, and points of subjective equality between transitions were computed. We found that $\Delta E_{ab}^*/s$ in CIELAB was not suitable for describing the perceived speed of temporal color changes in full-room illumination, since, for example, two hue transitions with the same physical speed of change, in terms of $\Delta E_{ab}^*/s$, were not perceived as changing at the same speed. This result confirmed that CIELAB, designed to characterize spatial color differences, was insufficiently accurate to predict the perception of temporal color transitions in illumination. Our results demonstrated

The developments and updates of the thesis, together with the source codes related to the Ph.D. project is available at: <https://github.com/kongfundamentals/Modeling-the-Temporal-Behavior-of-Human-Color-Vision-for-Lighting-Applications>.

that more uniformity was achieved by improving available color spaces such as CIELAB and DKL simply by introducing scaling factors between the axes spanning the color space.

In **Chapter 4**, we extended the stimuli to understand better how to improve the temporal uniformity of CIELAB further. The stimuli were presented in a square of 4.3° visual angle surrounded by a 4000 K adaptation field, which is a similar viewing condition as for which CIELAB was designed. The observers viewed pairs of temporal color transitions which were presented sequentially and were asked to select the one that appeared to change faster. The results confirmed that also under these conditions, CIELAB was not temporally uniform. We presented preliminary attempts to improve the temporal uniformity for both CIELAB and cone-excitation spaces (i.e., LMS and DKL). Despite the different viewing conditions in the experiments of **Chapter 3** and **Chapter 4**, the results for optimizing the scaling factors were in line with each other, which suggested the validity of the framework to improve the temporal uniformity.

In **Chapter 5**, a multiple-session psychophysical experiment was carried out to compare the accuracy, precision, and efficiency of five commonly used psychophysical methods for measuring the detection threshold of chromatic flicker. The methods employed were: (1) the classical 1-up-1-down staircase method with a yes-no task, (2) the weighted 3-up-1-down staircase method with a 2AFC task, (3) the method of constant stimuli with a 2AFC task, (4) the method of adjustment without reference, and (5) the method of adjustment with reference. The chromatic flicker stimuli were temporally modulated with a square wave, with the mean chromaticity of the two colors either 2700K, 4000K or 6500K, and the temporal frequency 2 Hz or 4 Hz. We found that each method has its advantages and disadvantages. We decided to continue the follow-up investigations using the method of adjustment without a reference, as it is very efficient, it is quite accurate compared to the assumed ground truth threshold, and the variance of the threshold is comparable to those of the other methods.

In **Chapter 6**, the modulation amplitude at which isoluminant chromatic flicker became visible was measured for three observers using the method of adjustment. The isoluminant stimuli were created for each observer individually, based on a technique similar to heterochromatic flicker photometry. The chromatic flicker stimuli were sinusoidal modulations, defined in the CIE 1976 UCS (u' , v') chromaticity diagram. To determine the flicker visibility threshold we varied the amplitude of modulations along four modulation directions around nine base colors and at seven different frequencies. The measured modulation amplitudes at the threshold were expressed as $1/\Delta(u', v')$, $1/\Delta LMS$ and $1/\Delta lms$, and the resulting thresholds were plotted as a function of frequency, resulting in 36 contrast sensitivity functions per observer, and were fitted with an exponential model. The model resulted in an average R²-value higher than 0.93 for the three different measures of contrast sensitivity. The two parameters of the model (i.e., the slope and intercept) were found to significantly depend on the base color and direction of the chromatic modulation. This means that $\Delta(u', v')$, ΔLMS , and Δlms are insufficiently suitable to predict the sensitivity to temporal chromatic modulations for different locations of the color space.

In **Chapter 7**, we collected data for the purpose of developing a temporal uniform color space taking into account individual differences in cone spectral sensitivities. Two psychophysical experiments were performed. In Experiment 1, we measured for fifteen participants the ratio in luminance between the two extreme colors of 60 chromatic flicker modulations at which the perceived flicker of that modulation was minimal. These so-called isoluminance ratios were used to estimate individual cone spectral sensitivities based on an

adapted version of the Individual Colorimetric Observer (ICO) model of Asano, Fairchild, and Blondé [5]. The ICO model described the perceived luminance better than the cone spectral sensitivities of the standard observer. In Experiment 2, we measured the temporal contrast sensitivity at fifteen base colors, four modulation directions, and three temporal frequencies for three participants that had substantially different isoluminance ratios. The results gave a first impression on the size of individual differences in sensitivity to chromatic temporal modulations. In addition, we defined two quantitative measures, i.e., *circularity* and *homogeneity*, to describe the local and global uniformity of color spaces, respectively. In general, the DKL space was somewhat more uniform than the LMS space. The ICO-based cone fundamentals improved the temporal uniformity of the color spaces both locally and globally. However, contrary to our expectations, the $u'_{10^\circ}, v'_{10^\circ}$ color space had overall the best *circularity* and *homogeneity*.

Our dataset on speed perception and flicker visibility of temporally modulated chromatic light demonstrated the effect of base color, modulation direction, and modulation frequency on the temporal sensitivity and provided a first estimate of individual differences in this sensitivity. Furthermore, the data showed great potential to correct existing color spaces towards a more temporally uniform color space. Although these corrections did not yield a sufficiently uniform temporal color space yet, our modeling efforts have provided valuable insights to light designers into how hue and chroma have to be changed in order to control the perceived speed of colored light transitions. In addition, we recommend a local uniformity (i.e., *circularity*) and global uniformity (i.e., *homogeneity*) measure to be used by standardization bodies to compare the performance of different color spaces.

List of Publications

JOURNAL PUBLICATIONS

1. **Xiangzhen Kong**, Michael J. Murdoch, Ingrid Vogels, Dragan Sekulovski, and Ingrid Heynderickx (2019). *Perceived Speed of Changing Color in Chroma and Hue Directions in CIELAB*. Journal of the Optical Society of America A 36(6), 1022-1032.
2. **Xiangzhen Kong**, Minchen Wei, Michael J. Murdoch, Ingrid Vogels, and Ingrid Heynderickx (2020). *Assessing the temporal uniformity of CIELAB hue angle*. Journal of the Optical Society of America A 37(4), 521-528.
3. **Xiangzhen Kong**, Ingrid Vogels, Dragan Sekulovski, and Ingrid Heynderickx (2021). *Modelling Sensitivity to Chromatic Temporal Modulations Using Individual Cone Fundamentals*. Submitted to Lighting Research & Technology.

CONFERENCE PUBLICATIONS

1. **Xiangzhen Kong**, Mijael R. Bueno Pérez, Ingrid Vogels, Dragan Sekulovski, and Ingrid Heynderickx (2018). *Modelling Contrast Sensitivity for Chromatic Temporal Modulations*. 26th Color and Imaging Conference Final Program and Proceedings, 324-329. Color and Imaging Conference (CIC), Vancouver, Canada.

POSTERS

1. **Xiangzhen Kong**, Jeroen Veelenturf, Ingrid Vogels, Dragan Sekulovski, and Ingrid Heynderickx (2016). *Comparison of Staircase and Constant Stimuli Methods for Measuring Chromatic Flicker Perception*. Perception Day (Dag van de Perceptie) 2016. Radboud University, Nijmegen, the Netherlands.
2. Michael J. Murdoch, **Xiangzhen Kong**, Ingrid Vogels, Dragan Sekulovski, and Ingrid Heynderickx (2018). *Dynamic Color Perception: How Fast Are the Munsell Hues?* Munsell Centennial Color Symposium: Bridging Science, Art, & Industry - Massachusetts College of Art and Design (MassArt), Boston, United States.
3. **Xiangzhen Kong**, Michael J. Murdoch, Ingrid Vogels, Dragan Sekulovski, and Ingrid Heynderickx (2018). *Perceived Speed of Changing Color in Chroma and Hue Directions in CIELAB*. Perception Day (Dag van de Perceptie) 2018. Radboud University, Nijmegen, the Netherlands.

Curriculum Vitæ

Xiangzhen Kong was born on 10-12-1990 in Shandong, China.

He received his B.Eng. in Software Engineering from Department of Computer Science and Technology at Wuhan University of Technology, China, in 2012. Afterward, he continued a joint Master-Ph.D. program in Computer Applications Technology in the same department. After completing the Master's part of the program, he joined Human-Technology Interaction (HTI) Group at the Department of Industrial Engineering & Innovation Sciences as a Ph.D. student at Eindhoven University of Technology in the Netherlands, of which the results are presented in this dissertation.

During his Ph.D. program, he has visited different academic institutions. From August to December 2017, he was a visiting scholar at Munsell Color Science Laboratory, Rochester Institute of Technology, U.S.A. In 2019, he was employed as a research assistant at Department of Building Services Engineering at The Hong Kong Polytechnical University, Hong Kong S.A.R., China, from August to December. Afterward, he did an internship at Apple Inc. in California, U.S.A., from January to September 2020. During this whole Ph.D. program, he closely collaborated with colleagues of Signify (former Philips Lighting, the Netherlands). He has published peer-reviewed journal papers and presented his research work at various workshops and international conferences.

Since March 2021, he started to work as a postdoc at ESAT/Light & Lighting Laboratory at KU Leuven, Belgium.

Acknowledgements

Sir Isaac Newton found himself in Cambridge in 1665 during the Great Plague of London. While he escaped the plague affecting Cambridge, he was 'in the prime of' his age 'for invention & minded Mathematics & Philosophy' more than any time since. He completed the famous prism experiment and kicked off a new era about light and color in a striking way. I can only be jealous, and I definitely tried to stand on his shoulder to see further. Without his pioneering work in optics and color, I would probably end somewhere else in another parallel world if it exists. A more important question: did he foresee that he would be a giant in his words when he was 'working from home?' Oh man, working from home was really hard - but like many things in life, we get used to it and take it as a new norm. The transition of the mindset was challenging, but when I started to work in front of my laptop with my pajamas, I know it's getting real. I know that I must have buttered myself up by mentioning Sir Isaac Newton here!

In the last two years, I was asked many times about when I am going to graduate, and I found myself always in the almost-there mindset. And I thought I would write my acknowledgments in Eindhoven, or in Hong Kong, or in Cupertino, or somewhere in China. They are all important stops during my Ph.D. journey. What do they say about life? "Life was like a box of chocolates. You never know what you're gonna get." I never expected that I would write down this piece of acknowledgments in Gent, Belgium. I have seen some parts of the world, and I am lucky to have met many kind, knowledgeable and supportive persons who have walked along with me during the journey. Consider the above texts as a starter of my acknowledgments.

First, I would love to give a big shout-out to all my supervisors: Prof. Dr. Ingrid Heynderickx, Dr. Ingrid Vogels, and Dr. Dragan Sekulovski! This six-year Ph.D. life has been an amazing and very colorful journey! Having three supervisors is not easy, especially when they (sometimes) have different opinions. But their different views have always inspired me and sparked new ideas. And I have learned to look at things from different perspectives, theoretical, practical, and whether it is fun to do. We managed to maximize our common ground and work together to move closer towards the goal of the Ph.D. project. For me, this dissertation is like my newborn baby, and my supervisors are definitely his godfather/godmothers!

Dear Ingrid Heynderickx, thank you so much for always being so supportive. I cannot imagine how you have managed to have so many back-to-back meetings but still always responded to my emails within a day or two. My productivity always peaks whenever I read 'before 7 A.M.' in your email. Your email is always like an espresso - concise, brief, only ten times stronger. They have been serving as productivity boosts during my working hours. You've always supported me to explore and experience different opportunities in my career development, and you always tried to guide me to make a better decision. I always enjoyed

and benefited from listening to your suggestions. I'm grateful that you always nudge me to finish my Ph.D. 'earlier,' and in the meanwhile, thank you for being so patient when there was a delay. You have always been thinking together with me and point an efficient path to walk towards my goal. I am deeply indebted.

Dear Ingrid Vogels, you interviewed me back at the beginning of 2015. I don't remember what exactly we talked about, but I think that the Skype meeting was one of the most exciting calls. I never thought I would come to the Netherlands to do a Ph.D., but I am very proud that I managed to do it! Thank you for not shutting the door for me. You have always been the calming presence during my whole Ph.D. life. I always tell my friends that you are one of the most elegant women that I know in person. I guess your Yoga spirits have brought me a calming feeling. I remember many times having our meeting in the old IPO building. The window was open, and the breeze, the green, and the sunshine. So beautiful. I enjoyed our meetings - every time, it felt more like a pleasant conversation. You have this magic power with you that whenever I feel stuck somewhere, your smile and encouragement helped me relax. And, thank you for telling me it is okay when I felt myself not being so productive. I realized that a good spirit is key to walk out of the holes and move on!

Dear Dragan, dear Long Ge, you are more like a big brother who is so knowledgeable, intelligent and you always have so many inspiring ideas. Working together with you never gave me the feeling that I am receiving instructions. It's more like constructing a Lego art together - there are just indefinite ways of doing things. I never worried about not having a way of doing something - I was only concerned about which way to go? The only limitation there is time. You have brought so much joy to my Ph.D. work. I feel very connected because sometimes I wonder, maybe you are more Chinese than me since you know so much about tea? I enjoyed listening to you share your experience with China, which sometimes also made me feel less homesick. You are a hotpot lover, and it's a pity that we didn't do it more often. Another pity is that we didn't attend a conference together, as I heard from Ingrid that it was so fun to attend a conference with you. I got panicked when the first time I had the meeting with you - because I wasn't sure whether we were talking about color or not. "DKL? What are you talking about?" But I am "relieved" that I can follow you later - much better! Thank you for making my Ph.D. life 'a bittle more bearable!'

My deep gratitude must also go to Dr. Michael Murdoch and Dr. Minchen (Tommy) Wei, with whom I worked together on parts of my Ph.D. project. Dear Michael, it was a great pleasure to work together with you on this exciting topic of dynamic lighting! I'm grateful for my trip to the Munsell Color Science Laboratory at RIT, which has greatly broadened my vision in color science. I learned so much from your course 'Computational Vision Science' and gained more hands-on experience in psychophysics, programming, and experimental design. Four months was short, and we were on a tight schedule. Thank you for allowing me sufficient room to pilot the experiment and provides all the support for all the resources that I needed. I'm proud that we managed to work together on a very innovative experiment! Besides, I had lots of fun spending time with the kids, and thanks (Hi Meredith) for inviting me for a full Thanksgiving experience! Dear Tommy, thanks for providing the opportunity for me to visit the Color and Illumination Laboratory at PolyU. It has been a unique experience. I enjoyed the brainstorming sessions where we quickly decided on the research question that we wanted to tackle. I learned a lot working together with you on system calibration and academic writing. Thanks for pointing out the important literature that I should be aware of.

I would like to extend my great gratitude to the other members of the defense committee, Prof. Dr. Mark Fairchild, Prof. Dr. Wilbert IJzerman, Prof. Dr. Astrid Kappers, and Prof. Dr. Peter Hanselaer, for the detailed review and comments to improve this dissertation. Special thanks to Prof. Dr. Peter Hanselaer. Thank you for being very supportive in the last phase of my Ph.D. journey and allows me sufficient time to wrap up my dissertation.

WHUT (my alma mater) & WHU

I gratefully acknowledge my supervisor Prof. Dr. Xiong Shengwu, at Wuhan University of Technology for always encouraging me to go on different adventures and exploring different research directions in the academic world. I had a fun time going on-site to collect data and work on projects that I was very interested in. I am deeply indebted to your supervision, guidance, and all kinds of support. I never thought about doing a Ph.D., and you are the one who introduces me to be on this academic path. Thank you for putting so much trust in me, although sometimes I failed your expectations (sorry!). My first trip abroad was to present our work at a conference in the USA, which has deeply encouraged me to contribute my ideas and efforts in the research world. I will not forget that you encouraged us to be open-minded and go to conferences to exchange ideas. I personally greatly benefited from it. Thanks for building the bridge for me to be exposed to the academic seminars at Wuhan University with Prof. Dr. Fang Zhixiang. I gained my first experience of interdisciplinary collaborating on research work. I am grateful to all the staff who helped me with the application for the CSC scholarship, and I am truly grateful for CSC providing me this great opportunity. Special thanks to Li Ling for reviewing my application materials. Thanks for Li Xiaobei, Li Jue, Ye Qiongyao, and Cheng Jianghong's help with all the paperwork. I'm also grateful for the support from graduate school. I was lucky to have the company of Du Xin and Yang Mengshi from Wuhan University while getting the paperwork done.

TUE & HTI & NL

I never thought I would be a Ph.D. student so far away from home. I attended a Ph.D. workshop in Beijing in 2014. Mrs. Marleen van Heusden kindly arranged an in-person interview for me. I had a nice interview with Dr. Karen Ali (thanks for your time!). I always find it amazing how my life trajectory has changed because of one interview. Later on, with my only knowledge of color (RGB), I had another interview with Ingrid V. I remember that evening (a sunny afternoon in the Netherlands. Very much later, I was able to see the window that I saw in Skype in person) very well. My roommates Zheng Senwen and Zhu Chifeng, stopped using the Internet bandwidths for me to have a good internet connection. Thank you all for making it happen. Later on, Mrs. C.M. (Kara) de Rooy and Mrs. Ellen de Bree (our dear former colleague) has helped me with all the paperwork. I was with Guo Jia and two other students who were all going to start their Ph.D. journey like me on the same flight. I was honestly a bit disappointed when I arrived at Eindhoven station - I guess I pictured a lot of skyscrapers when I read that Eindhoven is the fifth-largest city in the Netherlands. But later on, I am very surprised at how much I enjoyed living in Eindhoven. Li Peipeng picked me up (thank you) at the station on my first day of arrival and walked me to my first home - the 'space box' in the campus, less than one hundred meters away from my office at IPO. I had a nice start with some nice neighbors, Wang Qinyu (and her boyfriend Yorick), Li

Wenshu, and Pei Yulong. We explored the city together and had nice walks and runs near Karpendonkse Plas. And space box was the first place where I started to live alone for the first time, and I was able to learn and sharpen my cooking skills.

Zhang Chao was the first HTI colleague that I met in person, and he was also my office mate for more than three years. We often went to the restaurants in the city center and came back to spend the evenings together in the office at IPO - and it was nice to have his company, during which we had many inspiring conversations. Alain was another officemate, and he was the walking TUE Wikipedia who knew everything going on at TUE, and he introduced many Dutch music and anecdotes to me. Jia Lixiu was my very close friend in the first year, and I was very sad that she visited HTI for only one year. We come from the same province in China and share a lot of things in common. We had some good laughs, talked about research, and (more importantly) made delicious baozi together! I miss that. Wang Huihui was a cool HTI colleague and friend, we cycled from Eindhoven to Maastricht together, and I will never forget that trip. Liu Caixia was like a sister, and she is always so optimistic and always helped me. I was always relaxed when she was around. I was lucky to be her neighbor at some point, and we spent many Chinese festivals together, which greatly reduced my homesickness. I also enjoyed the company of Roy, who helped us a lot and helped me to understand how a 'Dutch' decision was made. (Thank you, Caixia and Roy!) Alejandro and I had a lot of clicks - maybe because we both used many 'staircases?' He has helped me a lot with doing better research management, and he is one of the most organized persons I know. He always grabbed me from my seat and had a 'how-are-you' coffee break talk. We had fun in the bars, and I almost always ordered fresh mint tea (forgive me, my friend). Samantha joined the group later, and I enjoyed sitting with her outside of the benches in the sunshine to have our refreshing coffee breaks. We shared our frustrations and encouraged each other during our Ph.D. trip. We also had fun times outside of work, playing board games and Switch games together. Heleen has a pearl of unique wisdom, and she always looks at things from a brighter side. I always enjoyed talking with her, and every time I felt refreshed. (Sorry that I didn't read all the books that you recommended!) Elcin is my tree-planting buddy, and we spent much time focusing on solving difficult problems together in the empty office building on Dutch winter rainy days during the lockdown - they were great spiritual support, and I gained a lot of mental power. I shall never forget how fun it can be to be committed to working with full concentration. And thank you for sharing the delicious Baklava stuffed with pistachio - sugar makes people happy! Anita is another close colleague and friend with who I had many nice walks and talks. We started to know each other better very much later due to an unexpected encounter near Dommel. I will never forget the ice-breaking walk (literally!) that we had!) She always helps the things arranged well in our group, and I know I can always rely on her when I had questions. Renata is another go-to person whenever I had questions or requests to the group (thank you for all the nice conversations in the hallway!) Armin always reminds me of the cool headmaster Albus Dumbledore, and he was very knowledgeable and kind. He has brought my colleagues and me lots of fun. I also enjoyed being in the same team with him for our group activities. Yvonne is another bright star in our group, and I enjoyed talking with her and got inspired by her talks on lighting. My experience of being a teaching assistant for Raymond's course (Perception and Motor Control) has greatly broadened my knowledge about perception, which was a great learning opportunity for me. And I also enjoyed the company of Mariska, Alejandro, and Margot being there to help the students together. Raymond has a lot of knowledge

about perception, and I was inspired a lot. I am also very grateful for the technical support from our lab manager Martin Boschman and another colleague Twan Aarts. They spent a lot of time and effort in helping me build a customized device (official name 'flicker box,' with a nickname being 'Deeper Blue') for my research. I shall never forget it. I liked my offices at both IPO and Atlas (hi officemates Maaïke and Vaida, thanks for the nice office conversations and coffee breaks!). I miss the brown-bag meetings, afternoon drinks at Intermate, lunch at Paviljoen (Peter always knocks on the door and asks "Lunch?" thank you, Peter, for making us have a very regular diet), lovely walks on the campus, Christmas lunches (Chris, thank you always dressing up like a real Christmas man, and for your nice Christmas-lunch opening speeches), escape-city game, HTI outgoings, amazing HTI band at all kinds parties and celebrations. They are an important part of my Ph.D. life! Also, I shouldn't forget the students that I enjoyed working with - my flicker-buddies: Mijael, I knew you at the HTI introduction lunch meeting, and it was a great surprise that I became your 'supervisor' for your master thesis. We had so much fun together and I was happy to work on the colorful and 'enjoyable' experiments (flicker flicker flicker, more flicker please!) with you. Thanks for sharing your enthusiasm for salsa with me too - no, not for me, I have two left feet; Jeroen (Veelenturf), Steffen, and Rik (Rik, don't stop if you read here, more will come later.) I really enjoyed working together with you on flicker (one should share the joy with others, right? Who doesn't love flicker?)

Outside of HTI, I was so lucky to meet other friends who have always supported me in all means. My colleague and friend Sun Qi from another research group. She is very 'green', and she would always encourage me to use reusable bags for my sandwiches when we had lunch together. And her awareness of sustainability has greatly influenced me. We have watched documentaries about the earth together, which provided me with a different perspective of looking at us human beings on this planet. We encouraged each other when facing Ph.D. dilemmas, and we had many nice coffee breaks together. On weekends, we had many nice hikes in the woods nearby. And I always loved her healthy desserts (less sugar!) Because of Sun Qi, I also became friends with Song Jiankang, with who I also had a nice time together. He became my running coach (oh, did it only happen once or twice? Interval training was hard!), and he was always a cool bro who has been very chill. Thank you both for celebrating an unforgettable 30th birthday for me! Speaking of running, I would also like to thank Xing Haitao for running together in winter and thanks for the cool toolbox. I would like to thank Ping Cui. We had many nice talks during lunch at the Traverse building and many walks together on the campus. Those cheerful moments refreshed my mind so that I could always concentrate on my work better. She is very supportive and inspiring. Her boyfriend Liu Zheng is another very optimistic bro. I am always recharged with an optimistic spirit when spending time together with them. Wang Zhijun, Shawnee, Ding Ding, Cao Shuqin, and I formed another bubble. I am so proud that we managed to plan so many successful outgoings together, in Delft, Hengelo, Enschede, Deventer, and Eindhoven, while we are all physically so far away from each other. Those meet-ups, sharing our frustrations, cheering each other up, and applauding for our milestones in our Ph.D. trajectories are invaluable to me. I miss our hotpot sessions together, our Christmas trip to Germany, our first 10km Marathon together in Leiden. Thank you, my friends. I also enjoyed the company of Wouter for my first longest cycling trip around Eindhoven, which probably marked the start of my leisure cycling adventures.

I also would like to express my thanks to Xin Yuanyuan, who has been friends with me

since high school. Destiny has brought us together in Wuhan, Xi'an, Chengdu, and I was excited that you also spent two years of your Ph.D. in Groningen. We have so much in common, and I was lucky to know you since a teenager. Our trips together have definitely added so much fun to me. And thank you for asking me to be one of your paranymphs at your defense ceremony.

In addition, I also would like to thank all the colleagues and staff members who have been organizing PROOF courses at TU/e, from which I greatly benefited in flourishing my skills set. Signify (former Philips Lighting) was also an important part of my Ph.D. project. I always enjoyed my stay at Signify and had nice talks with Dragan, Gosia, Bianca, Marcel, Marc, Annemieke, and others. In Maastricht, I enjoyed being a loyal audience of Mastreechter Staar, and I enjoyed all the performances during Christmas! And every time I had nice talks with some of the choir members. Special thanks to Jack (and his wife Marianne) for all the nice talks every time!

RIT & MCSL & USA

My trip to RIT was the first big surprise during my Ph.D., and it has a profound effect that has become more clear later on. As a student who came from another discipline, I found learning everything from textbooks and free online courses about color are sometimes inefficient (you are simply getting lost easily). Thanks to Ingrid (V.)'s initiative, Ingrid (H.)'s YES, the help of Michael, Val (Valerie Hemink), and the invitation letter from Mark Fairchild, my trip was smoothly arranged. I had read so many papers with Mark's name on them, and I was super excited! Very much later, I attended Mark's talks on different occasions (CIC and other online webinars), and I am always inspired by his knowledgeable and humorous talks.

Michael picked me up from the airport and drove me to the home-sweet-home in Brighton. That was one of the best choices that I ever made, and it was so fun to live together with Brian, Diana, Doriel, Rafi, and Rebecca. Brian works at RIT, and we drove (well, he drove) to RIT together every day. All the trips were so fun, and I had so many opportunities to pick up some idioms or sometimes some fun trivia. "Time flies like an arrow, what's up next?" He asked me. I was thinking hard. "Fruit flies like a banana." he added. He is probably one of the most colorful persons that I've ever met. He is a professional magician, he teaches programming (with sign language!) at RIT, he built his own vacation house and adventure park in the hidden woods, and he tells funny jokes! Spending time together with him was never too much! Not to mention he also made delicious pizza! Diana was a typical mom figure. She was always so caring and asked me how my day was or how I felt. She also put notes for me to warn me to check whether the food I put in the fridge expired or not. Rafi brought me to acroyoga, silent disco, and some other dancing events, all new to me. And he made his only bubble recipe which was a function of air humidity, ambient temperature, wind level... Amazing nerdy stuff. Doriel loves pickles, and every day I saw him swallow a full can of pickles (or two cans?). He and Rebecca are huge fans of swing dances, and they always invited me to join them. A lot of times, we played board games together, hating each other and laughing together. And we also went to the shows performed by Avi, which was super amazing! I had so much fun living there, and every day I was energetic to go to work. Many thanks for bringing me to experience the colorful local events (Rochester Fringe Festival) and driving me to do groceries (oh, I miss Wegmans!).

All the people at MCSL were super friendly and always willing to help. Val helped me

with arranging all the paperwork, and I had a very smooth start. Xie Hao drove me to the Ontario Beach Park, the Dinosaur Barbeque, the High Falls the first weekend I was there. We had many nice talks about research, and later on, we became close friends. We met in Boston, Cupertino, and traveled to Los Angeles together. That made me feel the world is so tiny! Wang Mingming constantly updated me with what was going on on the campus, and we had many fun lunchtimes together at the buffet. Mingming, thank you and Li Han for a nice trip to Niagara Falls. Samuel was another visiting scholar at RIT, and we had a very fun time together (hi María, and two kiddos!) in making Tiramisu (and talking about research). I am happy that we met again in Boston at a conference the next year after leaving RIT. And now he is a full professor in Spain, and I attended one of the online webinars that he organized! Small world! Zhang Lili, thanks for showing me around the campus, and your color matching experiment in VR has been ranked by me Top 5 most annoying experiments. Jiang Fu, thanks for giving me a lot of support when I came back from RIT, and thank you for broadcasting that I only drink ginger ale (for only once in Vancouver?) when I go to bars (hey, I remembered that I tried Bloody Mary in Boston together with you okay?). And Anku, thanks for sharing delicious authentic Indian food with us! And Nargess and others, thanks for participating in my experiments and supporting my research! I equipped myself with new knowledge about color vision while auditing some lectures by Michael, Roy, Susan, and Mark. And I missed the Monday pizza time together with other fellow students! If I remember it correctly, potato pizza was the most popular one. I missed the scavenger hunt together with Xie Hao, Yongmin, Katie, Olivia, Josh, and Zhang Lili. And thanks, Dawei, for bringing a lot of fun to us as well. I also met Liu Ping, who is a smart, diligent, and resilient student. We had quite a nice time together (webinars and hotpot!) Thanks to Diana, Brian, and Ping, I started to get to know the deaf community for the first time.

PolyU & Hong Kong

My trip to PolyU was a truly unique experience, and I witnessed something very unpeaceful. It was not easy to stay sane and focused on work during that period. But we (lab friends!) supported each other and kept each other up-to-date. Unlike what I had to check the weather before I left home for work when I was in the Netherlands, I had to check the road conditions when I was there. The subway and bus were unexpectedly interrupted by the unpleasant violence. I enjoyed the accompany of my lab friends Bao Wenyu, Hu Yu, Li Yiqian, Chen Siyuan, and Wang Jun. We spent most of the time together, and had amusing conversations during lunch, and cheer each other up during dinner time. I am also very grateful to my roommate Wen Guohua, for being so easy-going and cheerful! I missed our fried chicken moments and hotpot moments (in the 2,5 meter squared room)! I enjoyed going out for hikes with Li Yiqian and Zhang Yu to the remote islands in Hong Kong to embrace the great nature there. During my visit to PolyU, I also met a colleague from TU/e (surprise) whose office was upstairs, Tang Pei, who works at PolyU as a postdoc. We had nice talks and supported each other during the difficult times.

During my 'escape-from-Hong-Kong,' I had a great time together with some of my family members and friends. Thanks to friends Zhu Chifeng, Zhang Yu, Ding Xiaoyi, Zeng Jie, Fang Jinwei and Gan Xinyun for your company and great hospitality in Shenzhen. I was also very glad to meet up with my university friends with who I have close contact - Liu Jiaqi (13) and Zhou Hailun. Special thanks to Zhou Hailun for taking care of me during my stay at his

apartment in Guangzhou. Also, thanks to Ru Taotao (our visiting scholar at HTI) for showing me around Guangzhou. I had a great time together with you in the same office when you were at HTI. I shall also not forget my cousin Wu Nan and her husband Zhang Guojie for taking really good care of me during my stay in Shenzhen. Also, thanks for the care from my cousin Kong Xiangliang and my sister-in-law Ding Caiming (Also hi to my nephew and little niece) in Dongguan for a great holiday together. While I traveled to Wuhan, I am grateful for the arrangement of the department at my alma mater, that I had the pleasure to meet the teachers and professors that I was very grateful for. I was also delighted to meet my old university friends, Mao Yimei, Liu Jiaqi (hi 13 again!), Tan Kaixiang, Li Jie, Zhu Zhixing, and Xing Wenhui. I also felt much cared for by my cousin Kong Su Juan (it was a pity my brother-in-law Zhang Lingwei was out on a business trip when I was visiting), and I was happy to have dinner together with my cousin and my niece. Later on, I also met my TU/e friend He Chenchen in Wuhan. Thank you for managing to meet up for such a short time.

Another unexpected trip was to Hangzhou. Thanks, Prof. Ming Ronnier Luo, for organizing an inspiring conference at Zhejiang University. And thanks for Zhu Yuechen's help in taking care of the registration (Yuechen, thank you also for your company in Boston and Vancouver. I enjoyed visiting the highlights of the cities with you!).

Cupertino

I was blessed to have the invaluable opportunity to do an internship at Apple in California. And my Cupertino trip was truly amazing (it could be better if not were because of COVID.) But as my limited experience in the industry, I must admit that I realized the power of the knowledge that I learned during my Ph.D., and I also recognized the room for improvement in my skill set. I am very grateful to Mehmet and Beau. Thank you for your time in interviewing me. And thank you for your guidance and support during my internship. I am also grateful to the knowledgeable colleagues who work there, and I want to thank Ozzy, Amir, Francisco, Jun, Harini, Shahram, Amanda, Michael, Laura, Tim, Bill, Alex, Teun, Guo Kaikai, Zhang Chi, Liu Peng, Jackson, and many other friendly and inspiring souls! Thanks for the inspiring meetings and entertaining virtual hangouts. Also, thanks to Stacy at Cultural Vistas and Gerilyn for helping with all the paperwork. I was also happy to have the company of Emitis and Comrun (Cameron, hey buddy), who were also Ph.D. interns, there. I missed our walks at Apple Park. Cameron, buddy, we were encouraging each other the whole time. I missed our gelato sessions. We had nice lunch and dinner together, and I enjoyed our trip to Big Sur! Thanks for the nice memories!

Outside of work, I was blessed with the intern bubble with which I had a lot of fun and felt less isolated during the lockdown. Big shouts to my roommie Juan (you are an annoying kid!), Will (thank you for being my running buddy, and it was a great way of relaxation for me; thank you for introducing Chick-fil-A for me, and thanks for many refreshing walks together), Adriano, Justin, Tom, Yang Guowei (Frank), Clemence, Anis, and Randy! Thanks for making the lockdown life livable! Special thanks to Xie Hao (Hi Hao again) for showing me around and helping me start smoothly. And thanks for the LA-Disney trip initiative, and I had a fantastic long weekend together with you and your friends Yang Shuhan and Lyu Linpeng (thanks for the farewell homemade barbeque meal!). Great thanks also go to Wang Mingming and Li Han (hi again) for their great mental support when I needed it. Also, thanks to Luo Lin for meeting up and exchange ideas and the latest updates of the color

world (we met in many conferences and webinars, again, tiny world!).

I cannot thank my friends Zhou Baichuan and Von enough for their great support and company. They brought me disinfectant spray and hand sanitizers (even toilet papers) at the beginning of the pandemic because I couldn't get them in the nearby supermarkets. Thanks for driving me to explore different places in San Francisco, Pillor Point Harbor, and Mountain View, and thank you for driving me to experience delicious local restaurants and snack bars! We met in Wuhan, Boston, and never thought we would meet in Cupertino. You are a true north star (you better believe it)! And because of you, I was lucky to meet Fu Qiaobin and Xu Tingting, and we had spent so much fun time together (I am curious about your new apartment!).

After my Cupertino chapter, I had to wait for a few days before traveling back to the Netherlands due to visa issues. Thanks to the hospitality of my dear *shixiong* Chen Zhong and his wife, Zhang Weixiang. I had a great time in New Orleans and enjoyed the boat tour at the City Park, the Swap Tour to see crocodiles, delicious homemade seafood, pancakes, and most importantly, your company. Thank you for your prayers, and I am very grateful to be in your thoughts.

During my stay in Gent, I also want to express my gratitude to my friends and colleagues Rik (I enjoyed working together with you on the HTI research projects, and I enjoyed the coffee breaks and so many fun talks at IPO, Atlas and HTC! Although you are younger than me, I still always get many valuable advice from you! And, I shall never forget the HTC-tomato soup - we should visit again!), Sas, Hang, Ching-wei, Ma Shining, Ma Jiefei, Li Jiaye, Karel, and others. Special thanks to Rik and Sas for helping with the whole relocation. Without your help, moving during lockdown would be a nightmare.

I cannot put an exhaustive list to mention everyone's name, but I want to extend my thanks to Wang Hongchen for being a good friend since middle school, and we have experienced quite some adventures together. Thank you for always being with me. Thanks for being my teacher in financial management so that I survived well with my Ph.D. scholarship. My another friend since high school, Ren Guozhi, thank you for always checking in with me. I also want to thank my university friend Liu Xiangdong, for making Beijing feel like one of my important bases. I am very grateful for the company of Jing Tao and Liu Sijing (Alice) every time whenever I transit in Beijing. Also, thanks to the company of Kong Xianghai and Wang Hanyi when I was there. I would like to express my gratitude to Zhao Yanfei (suancaiyu), one of the friends who I have met in different cities (Qingdao, Wuhan, Xi'an, and Beijing). Thank you for always thinking of me! I would also like to thank my Chinese teacher Lian Shumin, for always caring about me, and I am grateful that she sent me a package of masks when the pandemic started in Cupertino.

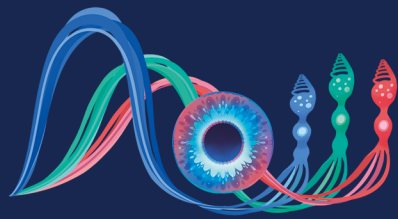
This acknowledgment is never complete without giving my deepest gratitude to my Chinese family and my Dutch family.

我心里有许多话，却发现用母语流畅地表达出来竟然如此困难，以致言不能传情，词无法达意。还记得上大学时第一次离家，只有寒暑假才回，同学中流行“从此，家乡只有冬夏，再无春秋”的说法，想不到而今冬夏也渐行渐远。时间如白驹过隙，常常未觉自己已到而立之年。时光倒回至二十几年前的夏天，我家平房的小院里，竹子茂密、无花果树繁盛、香椿本固枝荣，枣红色的方桌上盛满了家常饭菜。如今新楼拔地而起，小院的一花一木终成为照片上的风景，但这些温馨的记忆却历久弥新，它们会温暖地陪伴着我继续过好以后的人生。一路走来，我不会忘记家人无条件的支持，不会忘记与他们在欧亚大陆两端视频连线彼此分享的喜怒哀乐。感谢亲爱的

爸爸、妈妈、姐姐、姐夫，以及一直关心、爱护我的亲戚朋友和堂表兄弟姐妹们！你们永远是最坚实的后盾！“父兮生我，母兮鞠我。抚我畜我，长我育我，顾我复我，出入腹我。”谢谢妈妈，总是不厌其烦地叮嘱我在电脑前面坐久了要“起来走走”；谢谢爸爸，总是叮嘱我要踏实行事，谦虚有礼；谢谢姐姐，总是在乎我开不开心。小时候遥不可及的未来已来，我的人生也将开启全新的篇章。我知道我将继续带着你们的祝福与厚爱，迎接着未来的挑战与惊喜！

Jeroen, Iech höb diech gezag tot iech altied trök zal kaome nao mien reis naor Rochester, Hong Kong en Cupertino, mér dun aandere kant vaan de zaak is tots diech altied op miech höbs gewach. Bedaank dat's diech altied deej bezondere mins bis, dee altied um miech gief, good veur miech zörreg, en miech ech begriep. Iech höb ontelbaar veul prachtige herinneringe same met diech, aon de kleine paddestool zeuktoch in de bosse, in de lochballon bove Eindhoven, aon de fietstochte um de sjus gebore zwaone te bekieke, aon de wind bij Cabo da Roca, aon de oondergaonde zon op Texel, aon de stroumende rege onder de brögk in Warschau, aon de piramide vörmige Matterhorn, aon de zèlvere waterval vaan Lauterbrunnen, aon de rij veur ut bezeuk aon de Sint-Pietersbasiliek, aon de mauve blommezie bij Radio Kootwijk, aon de heuvels vaan Limburg, aon de kleure vaan de carnavalsoptoch Daanke totste miech zoe good ondersteuns höbs. Iech bin aoch erg gelökkig dat iech zoe geleefd bin door Magda, Hendrik en ama (iech hoop dat veer deze zomer eindelek weer same kinnen kaarten!) Lieve Magda, danke dats diech es un mam veur miech bis en vaan miech helds, um miech giefs en miech altied ondersteuns. Daanke veur ut neije vaan mien sokken, danke veur de walnoten deej veer altied vaan diech kriegen, danke dats diech miech altied aonmoedigs um gezoonder te eten. Lieve Hendrik, danke veur de heerleke diners deej stiech altied veur us kooks, bedaank dats diech us altied zoe good verzörgrs. Iech zal noets de mominten vergeten dat de witte asperges in mien moond smelten. Iech mis dien vers gekookte lasagne! Laote veer weer gaw samen geneten vaan un heerleke hotpot! Lieve ama, iech geneet altied vaan eeder bezeuk aon uuch, en iech wil uuch danken veur alle gezellige kaartspel mominten met get lekkers en appelsap tijdens de zonnige naomiddag vlaakbij Brusselse Poort (Daank uuch dat geer altied gooje kaarten veur miech afgoeis zoe dat Jeroen weer gefrustreerd kös zien, haha!) Muilkes!

Kong
June 2021
Gent, Belgium



Eindhoven University of Technology
Department of Industrial Engineering & Innovation Sciences